

BOUNDED REAL LEMMA AND STRUCTURED SINGULAR VALUE VERSUS DIAGONAL SCALING: THE FREE NONCOMMUTATIVE SETTING

JOSEPH A. BALL, GILBERT J. GROENEWALD, AND SANNE TER HORST

ABSTRACT. The structured singular value (often referred to simply as μ) was introduced independently by Doyle and Safanov as a tool for analyzing robustness of system stability and performance in the presence of structured uncertainty in the system parameters. While the structured singular value provides a necessary and sufficient criterion for robustness with respect to a structured ball of uncertainty, it is notoriously difficult to actually compute. The method of diagonal (or simply "D") scaling, on the other hand, provides an easily computable upper bound (which we call $\hat{\mu}$) for the structured singular value, but provides an exact evaluation of μ (or even a useful upper bound for μ) only in special cases. However it was discovered in the 1990s that a certain enhancement of the uncertainty structure (i.e., letting the uncertainty parameters be freely noncommuting linear operators on an infinite-dimensional separable Hilbert space) resulted in the D -scaling procedure leading to an exact evaluation of μ_{enhanced} ($\mu_{\text{enhanced}} = \hat{\mu}$), at least for the tractable special cases which were analyzed in complete detail. On the one hand this enhanced uncertainty has some appeal from the physical point of view: one can allow the uncertainty in the plant parameters to be time-varying, or more generally, one can catch the uncertainty caused by the designer's decision not to model the more complex (e.g. nonlinear) dynamics of the true plant. On the other hand, the precise mathematical formulation of this enhanced uncertainty structure makes contact with developments in the growing theory of analytic functions in freely noncommuting arguments and associated formal power series in freely noncommuting indeterminates. In this article we obtain the $\tilde{\mu} = \hat{\mu}$ theorem for a more satisfactory general setting.

1. INTRODUCTION

The structured singular value was introduced independently by Doyle [21] and Safanov [40]; see [45] for a thorough more recent treatment. Let N be a positive integer with a partitioning $N = n_1 + \cdots + n_s + m_1 + \cdots + m_f$ for positive integers n_i ($i = 1, \dots, s$) and m_j ($j = 1, \dots, f$). We let Δ denote the set of $N \times N$ matrices of the form

$$\Delta = \{\text{diag}[\delta_1 I_{n_1}, \dots, \delta_s I_{n_s}, \Delta_1, \dots, \Delta_f] : \delta_i \in \mathbb{C}, \Delta_j \in \mathbb{C}^{m_j \times m_j}\}. \quad (1.1)$$

For an $N \times N$ matrix $M \in \mathbb{C}^{N \times N}$, we define the *structured singular value* of M with respect to Δ by

$$\mu_{\Delta}(M) := \frac{1}{\min\{\|\Delta\| : \Delta \in \Delta, 1 \in \sigma(M\Delta)\}}, \quad (1.2)$$

2010 *Mathematics Subject Classification.* Primary 93D09; Secondary 93B28, 13F25, 47A60.

Key words and phrases. Structured singular value; Diagonal scaling; Free noncommutative function; Formal power series in free noncommuting indeterminates.

where in general $\sigma(X)$ denotes the spectrum of the square matrix X . Motivation for this notion comes from robust control theory (see [45, 22]).

In the case where $\mathbf{s} = 0$ and $\mathbf{f} = 1$, the structured singular value $\mu_{\Delta}(M)$ collapses to the largest singular value $\bar{\sigma}_1(M)$ of M or, equivalently, the induced operator norm of M as an operator on \mathbb{C}^N , where \mathbb{C}^N is given the standard 2-norm. A key property of the largest singular value from the point of view of systems and control follows from the Small Gain Theorem.

Theorem 1.1 (Small Gain Theorem). *Let $M \in \mathbb{C}^{N \times N}$ such that $\bar{\sigma}_1(M) < 1$. Then $I - \Delta M$ is invertible for all $\Delta \in \mathbb{C}^{N \times N}$ with $\|\Delta\| \leq 1$.*

The systems and control interpretation of this result is that $\bar{\sigma}_1(M) < 1$ implies that perturbation of the ‘plant’ M with a multiplicative perturbation Δ does not affect stability of the closed-loop feedback as long as $\|\Delta\| \leq 1$.

Another well known case is when $\mathbf{s} = 1$ and $\mathbf{f} = 0$. In that case $\mu_{\Delta}(M)$ coincides with the spectral radius of M , and hence $\mu_{\Delta}(M) < 1$ implies that $I - \delta M$ is invertible for all $\delta \in \mathbb{C}$ with $|\delta| \leq 1$.

There are many applications in which the uncertainty parameter Δ is known to carry some structure, as in (1.1). In these cases it is enough that the structured singular value $\mu_{\Delta}(M)$ be less than 1 to guarantee the maintenance of stability against structured multiplicative perturbations $\Delta \in \Delta$ with $\|\Delta\| \leq 1$.

However, it turns out that the structured singular value $\mu_{\Delta}(M)$ is notoriously difficult to compute in a computationally efficient and reliable way. Indeed, computing the exact structured singular value $\mu_{\Delta}(M)$ is an NP-hard problem [18].

There is a convenient upper bound for $\mu_{\Delta}(M)$ defined by

$$\hat{\mu}_{\Delta}(M) := \inf\{\|DM D^{-1}\| : D \in \Delta' \text{ and } D \text{ invertible}\},$$

where Δ' denotes the commutant of Δ in $\mathbb{C}^{N \times N}$, that is,

$$\Delta' = \{D \in \mathbb{C}^{N \times N} : D\Delta = \Delta D \text{ for all } \Delta \in \Delta\}. \quad (1.3)$$

It turns out that $\hat{\mu}_{\Delta}(M)$ can be computed accurately and efficiently. Indeed, to test whether $\hat{\mu}_{\Delta}(M) < 1$ it suffices to find a positive definite matrix $X \in \Delta'$ which solves the structured Stein inequality

$$M^* X M - X \prec 0.$$

Note that the condition $X \in \Delta'$ is equivalent to X having the block diagonal form

$$X = \text{diag}[X_1, \dots, X_{\mathbf{s}}, x_1 I_{m_1}, \dots, x_{\mathbf{f}} I_{m_{\mathbf{f}}}],$$

where X_i is a positive definite matrix of size $n_i \times n_i$ (for $i = 1, \dots, \mathbf{s}$) and x_j a positive number (for $j = 1, \dots, \mathbf{f}$). This puts the computation of $\hat{\mu}_{\Delta}$ within the framework of the MATLAB LMI toolbox.

While the general inequality $\mu_{\Delta}(M) \leq \hat{\mu}_{\Delta}(M)$ is easily derived, actual equality holds only in very special cases. In particular, equality holds for all M with respect to a given choice of structure specified by nonnegative integers \mathbf{s} and \mathbf{f} as in (1.1) if and only if $2\mathbf{s} + \mathbf{f} \leq 3$ (see [34, 45, 22]). Moreover, even with \mathbf{s} and \mathbf{f} in (1.1) fixed, there is in general no bound on the gap between $\mu_{\Delta}(M)$ and its upper bound $\hat{\mu}_{\Delta}(M)$; see [44]. Thus the compromise of using $\hat{\mu}_{\Delta}(M)$ as a substitute for $\mu_{\Delta}(M)$ can be arbitrarily conservative.

However, if the structure is relaxed by letting the uncertainty parameters δ_i and the matrix entries of Δ_j be operators on a separable infinite-dimensional Hilbert

space, say on $\ell^2 = \ell^2(\mathbb{Z}_+)$, the Hilbert space of square-summable complex sequences indexed by the nonnegative integers \mathbb{Z}_+ . Then the modified μ is equal to its easily computable upper bound. To make this precise, we introduce the enhanced structure

$$\tilde{\Delta} = \{\text{diag}[\tilde{\delta}_1 \otimes I_{\mathbb{C}^{n_1}}, \dots, \tilde{\delta}_s \otimes I_{\mathbb{C}^{n_s}}, \tilde{\Delta}_1, \dots, \tilde{\Delta}_f]\} \quad (1.4)$$

where each $\tilde{\delta}_i \in \mathcal{L}(\ell^2)$ and each $\tilde{\Delta}_j \in \mathcal{L}(\ell_{m_j}^2)$, with $\ell_{m_j}^2 = \mathbb{C}^{m_j} \otimes \ell^2$. We replace $M \in \mathbb{C}^{N \times N}$ with $\tilde{M} = M \otimes I_{\ell^2} \in \mathcal{L}(\ell_N^2)$ and define a new variation on $\mu(M)$ by

$$\tilde{\mu}_{\Delta}(M) := \mu_{\tilde{\Delta}}(I_{\ell^2} \otimes M).$$

It turns out that the two notions of $\hat{\mu}$ are the same:

$$\hat{\mu}_{\tilde{\Delta}}(I_{\ell^2} \otimes M) = \hat{\mu}_{\Delta}(M).$$

and hence the common value $\hat{\mu}_{\Delta}(M)$ is easily computable. The remarkable result is that this relaxed structured singular value is always equal to its easily computable upper bound, i.e.,

$$\tilde{\mu}_{\Delta}(M) = \hat{\mu}_{\Delta}(M). \quad (1.5)$$

This result can be found in the dissertation of Paganini [35] and is summarized in [32] without proof; the complete proof, as thoroughly elucidated in the book [22] (at least for the case where $s = 0$ with the case $s > 0$ indicated in the exercises) draws on earlier ideas and results from Megretski-Treil [33] and Shamma [41]. Also there is an interpretation of the quantity $\tilde{\mu}$ as robustness with respect to an enlarged block-structured uncertainty; one can view this enhanced block-structured uncertainty as allowing time-varying uncertainty in the system parameters, or, perhaps more appealingly, as specifying a range for the input-output pairs of the true plant, thus allowing for unmodeled dynamics (e.g. nonlinearities) in the behavior of the true plant (see [22, Chapter 8] for more complete details).

We mention that this result is but one more instance of a general phenomenon appearing often of late where a single-variable function theory result fails to have a compelling or complete generalization to the commutative multivariable setting, but does have a clean complete generalization to the free noncommutative setting; as for other examples, we mention the realization theory for rational matrix functions and for the Schur class on the unit disk (see [9, 10, 2]), Helton's result on representing a polynomial as a sum of squares [24], recent results in free noncommutative real algebraic geometry [19, 28], results on proper analytic maps [25, 26], as well as convexity theory [27, 29] and Nevanlinna-Pick interpolation [3].

As elegant as this result is, it is incomplete from a conceptual point of view since the structure given by (1.1) is limited in two respects:

- (L1) There is an asymmetry between the scalar blocks and the full blocks in (1.1). A scalar block $\delta_i I_{n_i}$ can be considered as a full block with size $m_i = 1$, but with a repetition (or multiplicity) of n_i possibly larger than 1 allowed. On the other hand, the full blocks Δ_j are considered to be independently arbitrary with no repetitions allowed.
- (L2) All blocks are considered square. There are interesting multidimensional input/state/output systems where this same structure occurs but with non-square blocks (see [9, 10]).

These limitations were addressed in the work of Ball-Groenewald-Malakorn [11] by making a connection with the earlier work of the same authors on the realization theory for so-called Structured Noncommutative Multidimensional Linear

Systems (SNMLSs), including a Kalman decomposition and state-space similarity theorem [9], together with a realization theorem for a noncommutative Schur-Agler class associated with conservative SNMLSs [10]. The structure of a SNMLS was encoded in an admissible graph \mathbf{G} , i.e., bipartite graph G carrying some additional structure together with a multiplicity function; see Section 3 for the precise setup. Motivation for introduction of this framework came from the quest for a more convenient coordinate-free way to analyze structures Δ as in (1.1) with the limitations (L1) and (L2) removed. The idea in [11] was to identify the resolvent expression $\Delta \mapsto (I - \Delta M)^{-1}$ as an element of the associated Schur-Agler class $\mathcal{SA}_{\mathbf{G}}(\mathcal{U})$ in case $\mu_{\mathbf{G}}(M) < 1$. However, this identification required an unnecessary additional hypothesis making the analysis in [11] incomplete. One of the contributions of the present paper is to adapt one piece of the analysis in [35, 22] to verify a key lemma (see Lemma 4.3) which implies that this additional hypothesis indeed can be removed and thereby to complete the analysis begun in [11].

A second contribution of the present paper is to identify the extra ingredient needed to show how the techniques of Dullerud-Paganini [35, 22] can be adapted to get (1.5) in full generality (without the limitations (L1) and (L2)); the precise result is formulated in our Main Result (Theorem 3.2).

We also show how our Main Result itself can be used to get an alternative proof of the realization theorem for the noncommutative Schur-Agler class $\mathcal{SA}_{\mathbf{G}}(\mathcal{U}, \mathcal{V})$, at least for the finite-dimensional case (see Remark 4.4); thus one can argue that the main result of [10] was already implicitly contained in the 1996 dissertation of Paganini [35]. It is interesting to note that the proof based on [11] requires the realization theorem for the noncommutative Schur-Agler class $\mathcal{SA}_{\mathbf{G}}(\mathcal{U}, \mathcal{V})$ which ultimately relies on an infinite-dimensional cone-separation argument, while the proof of Dullerud-Paganini [35, 22] uses a more elementary finite-dimensional cone-separation argument. We should also mention that relatively recent results of K ro glu-Scherer [31] also remove the limitations (L1) and (L2) and present still finer results concerning robust stability/performance against a fine class of structured uncertainties Δ (see Remark 5.7 below).

The paper is organized as follows. Section 2 reviews notation and results concerning tensor product spaces which will be needed in the sequel; this includes an adaptation of the Douglas lemma to the higher multiplicity setup, which is the extra ingredient needed to carry out the Dullerud-Paganini proof of the Main Result for the higher multiplicity situation. In Section 3 we recall the graph formalism from [9, 10, 11] and reformulate the desired result (1.5) in this framework for the general setting. In Section 4 we identify and prove the key lemma needed to complete the analysis from [11] and thereby get our first proof of the Main Result, Theorem 3.2 below. In Section 5 we show how the analysis of Dullerud-Paganini can be beefed up to handle the more general case with limitations (L1) and (L2) removed. In Section 6 we show how the alternative enhanced structured singular value of Bercovici-Foias-Khargonekar-Tannenbaum [17] can be handled by the same type of convexity analysis as used by Dullerud-Paganini.

A preliminary version of this report was given in the conference proceedings paper [12].

2. PRELIMINARIES ON TENSOR PRODUCTS

Let \mathcal{H} and \mathcal{K} be two Hilbert spaces. We shall have use for a fixed conjugation operator \mathcal{C} on \mathcal{K} , i.e., an operator \mathcal{C} on \mathcal{K} with the following properties:

- (i) $\mathcal{C}(\alpha f + g) = \bar{\alpha}\mathcal{C}(f) + \mathcal{C}(g)$ (anti-linear)
- (ii) $\langle \mathcal{C}f, \mathcal{C}g \rangle = \langle g, f \rangle = \overline{\langle f, g \rangle}$ (isometric)
- (iii) $\mathcal{C}^2 = I$ (involution)

To construct such an operator, choose any orthonormal basis $\{e_j : j \in A\}$ for \mathcal{K} and define \mathcal{C} by

$$\mathcal{C} : \sum_{j \in A} c_j e_j \mapsto \sum_{j \in A} \bar{c}_j e_j$$

where \bar{c}_j is the ordinary complex conjugate of the complex number c_j . For convenience of notation we shall often write \bar{k} instead of $\mathcal{C}k$.

The Hilbert space tensor product $\mathcal{H} \otimes \mathcal{K}$ is defined as the completion of the linear span of the pure tensor elements $h \otimes k$ where the inner product on pure tensors is given by

$$\langle h \otimes k, h' \otimes k' \rangle_{\mathcal{H} \otimes \mathcal{K}} = \langle h, h' \rangle_{\mathcal{H}} \langle k, k' \rangle_{\mathcal{K}}.$$

We note that in this construction the pure tensor $ch \otimes k$ is identified with the pure tensor $h \otimes ck$ for $c \in \mathbb{C}$ a scalar. It is convenient to view a vector h in the Hilbert space \mathcal{H} also as an operator $h \in \mathcal{L}(\mathbb{C}, \mathcal{H})$:

$$h : c \mapsto c \cdot h \in \mathcal{H} \text{ for } c \in \mathbb{C}.$$

with adjoint $h^* : \mathcal{H} \rightarrow \mathbb{C}$ given by

$$h^* : h' \mapsto \langle h, h' \rangle_{\mathcal{H}} \in \mathbb{C}.$$

With this interpretation, the Hilbert space inner product itself can be rewritten as

$$\langle h, h' \rangle_{\mathcal{H}} = (h')^* h.$$

A space closely related to the Hilbert space tensor product $\mathcal{H} \otimes \mathcal{K}$ is the space $\mathcal{C}_2(\mathcal{K}, \mathcal{H})$ of Hilbert-Schmidt operators from \mathcal{H} into \mathcal{K} , i.e., the space of operators $T \in \mathcal{L}(\mathcal{K}, \mathcal{H})$ such that T^*T is in the trace class $\mathcal{C}_1(\mathcal{K}) = \mathcal{C}_1(\mathcal{K}, \mathcal{K})$. These operators form a Hilbert space with inner product given by

$$\langle S, T \rangle_{\mathcal{C}_2(\mathcal{K}, \mathcal{H})} = \text{tr}(T^*S).$$

In fact, the following result gives a useful identification between the tensor-product Hilbert space $\mathcal{H} \otimes \mathcal{K}$ and the Hilbert space of Hilbert-Schmidt operators $\mathcal{C}_2(\mathcal{K}, \mathcal{H})$. For completeness we include an elementary proof; a good reference for more general tensor-product constructions is the book of Takesaki [43].

Proposition 2.1. *Let \mathcal{H} and \mathcal{K} be two Hilbert spaces with a fixed conjugation operator $\mathcal{C} : k \mapsto \bar{k}$ given on \mathcal{K} . Define a map $U_{\mathcal{H}, \mathcal{K}}$ on pure tensors in $\mathcal{H} \otimes \mathcal{K}$ into rank-1 operators from \mathcal{K} into \mathcal{H} according to the formula*

$$U_{\mathcal{H}, \mathcal{K}} : h \otimes k \mapsto h(\bar{k})^* =: hk^\top.$$

Then $U_{\mathcal{H}, \mathcal{K}}$ extends by linearity and continuity to a unitary map from the Hilbert space $\mathcal{H} \otimes \mathcal{K}$ onto the Hilbert space $\mathcal{C}_2(\mathcal{K}, \mathcal{H})$.

Proof. For purposes of the proof, we abbreviate $U_{\mathcal{H}, \mathcal{K}}$ to U . As $\mathcal{H} \otimes \mathcal{K}$ is the Hilbert space completion of the span of the pure tensors and $\mathcal{C}_2(\mathcal{K}, \mathcal{H})$ is the Hilbert space completion of the span of the rank-one operators, it suffices to check that U preserves the respective inner products on pure tensors:

$$\langle U[h \otimes k], U[h' \otimes k'] \rangle_{\mathcal{C}_2(\mathcal{K}, \mathcal{H})} = \langle h \otimes k, h' \otimes k' \rangle_{\mathcal{C}_2(\mathcal{K}, \mathcal{H})}. \quad (2.1)$$

To this end, we compute

$$\begin{aligned} \langle U[h \otimes k], U[h' \otimes k'] \rangle_{\mathcal{C}_2(\mathcal{K}, \mathcal{H})} &= \langle h\bar{k}^*, h'(\bar{k}')^* \rangle_{\mathcal{C}_2(\mathcal{K}, \mathcal{H})} = \text{tr}(\bar{k}'(h')^* h\bar{k}^*) \\ &= \text{tr}((h')^* h\bar{k}^* \bar{k}') = \langle h, h' \rangle_{\mathcal{H}} \cdot \langle \bar{k}', \bar{k} \rangle_{\mathcal{K}} = \langle h, h' \rangle_{\mathcal{H}} \cdot \langle k, k' \rangle_{\mathcal{K}} \\ &= \langle h \otimes k, h' \otimes k' \rangle_{\mathcal{H} \otimes \mathcal{K}} \end{aligned}$$

as required. \square

Given four Hilbert spaces $\mathcal{H}, \mathcal{K}, \mathcal{H}_0, \mathcal{K}_0$ and operators $X \in \mathcal{L}(\mathcal{H}, \mathcal{K})$ and $Y \in \mathcal{L}(\mathcal{H}_0, \mathcal{K}_0)$, the tensor-product operator $X \otimes Y$ is defined on pure tensors in $\mathcal{H} \otimes \mathcal{H}_0$ according to the formula

$$X \otimes Y : h \otimes h_0 \mapsto Xh \otimes Yh_0 \in \mathcal{K} \otimes \mathcal{K}_0. \quad (2.2)$$

It is not hard to see that $X \otimes Y$ extends to a bounded operator from $\mathcal{H} \otimes \mathcal{H}_0$ into $\mathcal{K} \otimes \mathcal{K}_0$ with $\|X \otimes Y\|_{\mathcal{L}(\mathcal{H} \otimes \mathcal{H}_0)} = \|X\|_{\mathcal{L}(\mathcal{H}, \mathcal{K})} \cdot \|Y\|_{\mathcal{L}(\mathcal{H}_0, \mathcal{K}_0)}$. A convenient tool for working with such operators is to use the identification maps $U_{\mathcal{H}, \mathcal{K}}$ and $U_{\mathcal{H}_0, \mathcal{K}_0}$ to view $X \otimes Y$ as acting between Hilbert-Schmidt operator spaces $\mathcal{C}_2(\mathcal{H}_0, \mathcal{H})$ and $\mathcal{C}_2(\mathcal{K}_0, \mathcal{K})$ instead; indeed this is one approach to seeing why $X \otimes Y$ is bounded with norm as in (2.2). Here we use the notation Y^\top for the operator

$$Y^\top : k \mapsto \overline{Y^* k}.$$

Proposition 2.2. *Given $X \in \mathcal{L}(\mathcal{H}, \mathcal{K})$ and $Y \in \mathcal{L}(\mathcal{H}_0, \mathcal{K}_0)$, let L_X be the left multiplication operator $L_X : T \mapsto XT$ mapping the Hilbert-Schmidt-operator space $\mathcal{C}_2(\mathcal{K}_0, \mathcal{H})$ to the Hilbert-Schmidt-operator space $\mathcal{C}_2(\mathcal{K}_0, \mathcal{K})$, and let R_{Y^\top} be the right multiplication operator $R_Y : T' \mapsto T'Y^\top$ mapping the Hilbert-Schmidt-operator space $\mathcal{C}_2(\mathcal{H}_0, \mathcal{H})$ to the Hilbert-Schmidt-operator space $\mathcal{C}_2(\mathcal{K}_0, \mathcal{H})$. If $U_{\mathcal{H}, \mathcal{H}_0} : \mathcal{H} \otimes \mathcal{H}_0 \rightarrow \mathcal{C}_2(\mathcal{H}_0, \mathcal{H})$ and $U_{\mathcal{K}, \mathcal{K}_0} : \mathcal{K} \otimes \mathcal{K}_0 \rightarrow \mathcal{C}_2(\mathcal{K}_0, \mathcal{K})$ are the identification maps as introduced in Proposition 2.1, then we have the intertwining relation*

$$U_{\mathcal{K}, \mathcal{K}_0}(X \otimes Y) = L_X R_{Y^\top} U_{\mathcal{H}, \mathcal{H}_0}. \quad (2.3)$$

Proof. It suffices to verify that the relation (2.3) holds when applied to an elementary tensor $h \otimes h_0$. We compute

$$\begin{aligned} U_{\mathcal{K}, \mathcal{K}_0}[(X \otimes Y)(h \otimes h_0)] &= U_{\mathcal{K}, \mathcal{K}_0}[Xh \otimes Yh_0] = (Xh)(\overline{Yh_0})^* = (Xh)(\overline{h_0}^* Y^\top) \\ &= X(h\overline{h_0}^*)Y^\top = L_X R_{Y^\top} U_{\mathcal{H}, \mathcal{H}_0}[h \otimes h_0] \end{aligned}$$

as required. \square

The well-known Douglas lemma (see [20]) asserts that, given Hilbert space operators $A \in \mathcal{L}(\mathcal{Y}, \mathcal{Z})$ and $B \in \mathcal{L}(\mathcal{X}, \mathcal{Z})$, there exists an operator $X \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ with $AX = B$ and $\|X\| \leq 1$ if and only if $BB^* - AA^* \preceq 0$. We shall have use of the adjoint version: *given Hilbert space operators $A \in \mathcal{L}(\mathcal{Z}, \mathcal{Y})$ and $B \in \mathcal{L}(\mathcal{Z}, \mathcal{X})$, then there exists an operator $X \in \mathcal{L}(\mathcal{Y}, \mathcal{Z})$ satisfying $XA = B$ with $\|X\| \leq 1$ if and only if $B^*B - A^*A \preceq 0$.* The special case where $\mathcal{Z} = \mathbb{C}$ appears as Lemma 8.4 in [22] and is crucial for the proof of the multiplicity-one special case of Theorem 3.2 there.

The following structured version of the Douglas lemma is crucial for the second proof of our main result, Theorem 3.2, for the general case.

Proposition 2.3. *Suppose that we are given three Hilbert spaces \mathcal{H} , \mathcal{K} , \mathcal{H}_0 , along with vectors $p \in \mathcal{H} \otimes \mathcal{H}_0$ and $q \in \mathcal{K} \otimes \mathcal{H}_0$. Then there exists an operator $X \in \mathcal{L}(\mathcal{H}, \mathcal{K})$ satisfying*

$$(X \otimes I_{\mathcal{H}_0})p = q \text{ and } \|X\|_{\mathcal{L}(\mathcal{H}, \mathcal{K})} \leq 1 \quad (2.4)$$

if and only if

$$U_{\mathcal{K}, \mathcal{H}_0}[q]^* U_{\mathcal{K}, \mathcal{H}_0}[q] - U_{\mathcal{H}, \mathcal{H}_0}[p]^* U_{\mathcal{H}, \mathcal{H}_0}[p] \preceq 0. \quad (2.5)$$

Proof. Application of the identification maps $U_{\mathcal{H}, \mathcal{H}_0}$ and $U_{\mathcal{K}, \mathcal{H}_0}$ combined with the intertwining relation (2.3) given by Proposition 2.2 transforms the problem of finding an X satisfying (2.4) to: *find $X \in \mathcal{L}(\mathcal{H}, \mathcal{K})$ with $\|X\|_{\mathcal{L}(\mathcal{H}, \mathcal{K})} \leq 1$ so that*

$$X U_{\mathcal{H}, \mathcal{H}_0}[p] = U_{\mathcal{K}, \mathcal{H}_0}[q].$$

The criterion for a solution of this problem is then given by the standard Douglas lemma (in adjoint form) resulting in (2.5) as the criterion for the existence of a solution. \square

3. THE GRAPH FORMALISM

For the remainder of this paper let $G = (\mathbf{V}, \mathbf{E})$ be a finite simple undirected bipartite graph such that each path-connected component of G is a complete bipartite graph. Here \mathbf{V} denotes the set of vertices and \mathbf{E} the set of edges. Since G is a bipartite graph, the vertex set \mathbf{V} admits a decomposition $\mathbf{V} = \mathbf{S} \cup \mathbf{R}$, with $\mathbf{S} \cap \mathbf{R} = \emptyset$, such that each edge $e \in \mathbf{E}$ has one vertex in \mathbf{S} (the source side), denoted by $s(e)$, and one vertex in \mathbf{R} (the range side), denoted by $r(e)$. We let \mathbf{P} denote the set of path-connected components of G . We let \mathbf{P} denote the set of path-connected components of G . For a vertex $v \in \mathbf{V}$ we let $[v]$ indicate the path-connected component $p \in \mathbf{P}$ that contains v . For each $p \in \mathbf{P}$ we denote the vertex set and edge set of p by \mathbf{V}_p and \mathbf{E}_p , respectively. Each path-connected component p of G is also a simple bipartite graph and its vertex set \mathbf{V}_p can be decomposed as $\mathbf{V}_p = \mathbf{S}_p \cup \mathbf{R}_p$ with $\mathbf{S}_p = \mathbf{S} \cap \mathbf{V}_p$ and $\mathbf{R}_p = \mathbf{R} \cap \mathbf{V}_p$. By the assumption that each path-connected component is a complete bipartite graph, for each $p \in \mathbf{P}$, the set \mathbf{E}_p consists of all possible edges connecting a vertex in \mathbf{S}_p with a vertex in \mathbf{R}_p . By definition of connected component, no edge e of G connects a vertex in $\mathbf{S}_p \cup \mathbf{R}_p$ with a vertex in $\mathbf{S}_{p'} \cup \mathbf{R}_{p'}$ if $p \neq p'$.

We shall on occasion want also to specify a *multiplicity structure* to such a graph G ; by this we mean a specification of a Hilbert space \mathcal{H}_p for each path-connected component $p \in \mathbf{P}$ of G . We then say that the whole collection $\mathbf{G} = (G, \{\mathcal{H}_p : p \in \mathbf{P}\})$ is an *admissible graph with multiplicity*, or an *M-graph* for short. Finally, for the most general version of the structure, we will specify a C^* -algebra Δ_p represented concretely as a C^* -subalgebra of $\mathcal{L}(\mathcal{H}_p)$; we call this more elaborate structure $\overline{\mathbf{G}} = (G, \{\Delta_p \subset \mathcal{L}(\mathcal{H}_p)\})$ a *admissible graph with specified C^* -algebras*, or *A-graph* for short.

3.1. The uncertainty structure: general case. Let $\overline{\mathbf{G}} = (G, \{\Delta_p \subset \mathcal{L}(\mathcal{H}_p)\})$ be an A-graph as defined above. We set $\mathcal{H}_v = \mathcal{H}_{[v]}$ for each $v \in V$ and we further

introduce the spaces

$$\begin{aligned}\mathcal{H}_{\mathbf{S}} &= \bigoplus_{s \in \mathbf{S}} \mathcal{H}_s, & \mathcal{H}_{\mathbf{S},p} &= \bigoplus_{s \in \mathbf{S}_p} \mathcal{H}_s \quad (p \in \mathbf{P}), \\ \mathcal{H}_{\mathbf{R}} &= \bigoplus_{r \in \mathbf{R}} \mathcal{H}_r, & \mathcal{H}_{\mathbf{R},p} &= \bigoplus_{r \in \mathbf{R}_p} \mathcal{H}_r \quad (p \in \mathbf{P}).\end{aligned}\tag{3.1}$$

For $s \in \mathbf{S}$ we write ι_s for the canonical embedding of $\mathcal{H}_{[s]}$ into $\mathcal{H}_{\mathbf{S}}$ that identifies $\mathcal{H}_{[s]}$ with the s -th component $\mathcal{H}_s = \mathcal{H}_{[s]}$ in the direct sum defining $\mathcal{H}_{\mathbf{S}}$ in (3.1): $\iota_s h = \bigoplus_{s' \in \mathbf{S}} (\delta_{s',s} h)$ for $h \in \mathcal{H}_{[s]}$, with $\delta_{s',s}$ equal to the Kronecker delta. Similarly, for $r \in \mathbf{R}$ we write ι_r for the embedding of $\mathcal{H}_{[r]}$ as the r -th component $\mathcal{H}_r = \mathcal{H}_{[r]}$ in the direct-sum defining $\mathcal{H}_{\mathbf{R}}$ in (3.1). Note that ι_s (respectively ι_r) acts on $\mathcal{H}_{[s]}$ (respectively $\mathcal{H}_{[r]}$) and not on \mathcal{H}_s (respectively \mathcal{H}_r), so that for an $e \in E$ the product $\iota_{s(e)} \iota_{r(e)}^*$ is properly defined.

We let $\Delta^{\mathbf{E}}$ denote the set of all operator-tuples $Z = (Z_e)_{e \in \mathbf{E}}$ indexed by the edge set \mathbf{E} such that the component Z_e is in the C^* -algebra $\Delta_{[\mathbf{r}(e)]} = \Delta_{[\mathbf{s}(e)]}$. Given any $Z = (Z_e)_{e \in \mathbf{E}} \in \Delta^{\mathbf{E}}$, we define an operator $L_{\overline{\mathbf{G}}}(Z) \in \mathcal{L}(\mathcal{H}_{\mathbf{R}}, \mathcal{H}_{\mathbf{S}})$ by

$$L_{\overline{\mathbf{G}}}(Z) = \sum_{e \in \mathbf{E}} \iota_{\mathbf{s}(e)} Z_e \iota_{\mathbf{r}(e)}^*. \tag{3.2}$$

We then define the uncertainty set $\Delta_{\overline{\mathbf{G}}}$ associated with the A -graph $\overline{\mathbf{G}}$ by

$$\Delta_{\overline{\mathbf{G}}} = \{L_{\overline{\mathbf{G}}}(Z) : Z = (Z_e)_{e \in \mathbf{E}} \in \Delta^{\mathbf{E}}\} \subset \mathcal{L}(\mathcal{H}_{\mathbf{R}}, \mathcal{H}_{\mathbf{S}}). \tag{3.3}$$

Since the elements of $\Delta_{\overline{\mathbf{G}}}$ in general are not square, we cannot work with its commutant, like we did with Δ in (1.3). Instead we will make use of the intertwining space

$$\Delta'_{\overline{\mathbf{G}}} = \{(X, Y) \in \mathcal{L}(\mathcal{H}_{\mathbf{R}}) \times \mathcal{L}(\mathcal{H}_{\mathbf{S}}) : \Delta X = Y \Delta, \Delta \in \Delta_{\overline{\mathbf{G}}}\}. \tag{3.4}$$

The following proposition gives an explicit description of this intertwining space.

Proposition 3.1. *The set $\Delta'_{\overline{\mathbf{G}}}$ is given by*

$$\Delta'_{\overline{\mathbf{G}}} = \{(X, Y) : X = \sum_{r \in \mathbf{R}} \iota_r \Gamma_{[r]} \iota_r^*, Y = \sum_{s \in \mathbf{S}} \iota_s \Gamma_{[s]} \iota_s^* \text{ where } \Gamma_p \in \Delta'_p, p \in \mathbf{P}\}.$$

Here Δ'_p denotes the commutant of Δ_p in $\mathcal{L}(\mathcal{H}_p)$.

Proof. Assume the C^* -algebras Δ_p , $p \in P$, are unital. If this is not the case then one can modify the argument using approximate identities. Let $(X, Y) \in \Delta'_{\overline{\mathbf{G}}}$. Choose an $e_0 \in \mathbf{E}$ and take $Z_{e_0} = I$ and $Z_{e'} = 0$ for all $e' \neq e_0$. With this choice of $Z = (Z_e)_{e \in \mathbf{E}} \in \Delta^{\mathbf{E}}$ the intertwining relation $L_{\overline{\mathbf{G}}}(Z)X = Y L_{\overline{\mathbf{G}}}(Z)$ yields

$$\iota_{\mathbf{s}(e_0)} \iota_{\mathbf{r}(e_0)}^* X = Y \iota_{\mathbf{s}(e_0)} \iota_{\mathbf{r}(e_0)}^*.$$

Since $\iota_v^* \iota_v = I$ and $\iota_v^* \iota_{v'} = 0$ for all $v, v' \in V$ with $v \neq v'$ (and v and v' either both in \mathbf{S} or both in \mathbf{R}), we have

$$\iota_{\mathbf{r}(e_0)}^* X \iota_{\mathbf{r}(e_0)} = \iota_{\mathbf{s}(e_0)}^* Y \iota_{\mathbf{s}(e_0)}, \quad \iota_{\mathbf{r}(e_0)}^* X \iota_r = 0 \ (r \neq \mathbf{r}(e_0)), \quad \iota_s^* Y \iota_{\mathbf{s}(e_0)} = 0 \ (s \neq \mathbf{s}(e_0)).$$

Set $X_r = \iota_r^* X \iota_r$ and $Y_s = \iota_s^* Y \iota_s$ for each $r \in \mathbf{R}$ and each $s \in \mathbf{S}$. Since $e_0 \in \mathbf{E}$ was chosen arbitrarily, the above identities imply that

$$X = \sum_{r, r' \in \mathbf{R}} \iota_r \iota_{r'}^* X \iota_{r'} \iota_r^* = \sum_{r \in \mathbf{R}} \iota_r \iota_r^* X \iota_r \iota_r^* = \sum_{r \in \mathbf{R}} \iota_r X_r \iota_r^*$$

and similarly

$$Y = \sum_{s,s' \in \mathbf{S}} \iota_s \iota_s^* Y \iota_{s'} \iota_{s'}^* = \sum_{s \in \mathbf{S}} \iota_s \iota_s^* Y \iota_s \iota_s^* = \sum_{s \in \mathbf{S}} \iota_s Y \iota_s^*.$$

Furthermore

$$X_r = X_{r'} = Y_s = Y_{s'} \text{ whenever } [r] = [r'] = [s] = [s'].$$

We conclude that there is a well-defined operator Γ_p on \mathcal{H}_p given by

$$\Gamma_p = X_r = Y_s \text{ whenever } [r] = [s] = p$$

and that X and Y are given by

$$X = \sum_{r \in \mathbf{R}} \iota_r \Gamma_{[r]} \iota_r^*, \quad Y = \sum_{s \in \mathbf{S}} \iota_s \Gamma_{[s]} \iota_s^*. \quad (3.5)$$

We show next that $\Gamma_p \in \Delta'_p$ for each p . Indeed, fix a $p \in \mathbf{P}$ choose $\Delta_p \in \Delta_p$ and let $e_0 \in \mathbf{E}$ such that $[s(e)] = p$. We take $Z = (Z_e)_{e \in E} \in \Delta^{\mathbf{E}}$ with $Z_{e_0} = \Delta_p$ and $Z_{e'} = 0$ for $e' \neq e_0$. Then $L_{\overline{\mathbf{G}}}(Z)X = YL_{\overline{\mathbf{G}}}(Z)$ yields

$$\begin{aligned} \iota_{s(e_0)} \Delta_p \Gamma_p \iota_{\mathbf{r}(e_0)}^* &= \iota_{s(e_0)} \Delta_p \iota_{\mathbf{r}(e_0)}^* \iota_{\mathbf{r}(e_0)} \Gamma_p \iota_{\mathbf{r}(e_0)}^* = \iota_{s(e_0)} \Delta_p \iota_{\mathbf{r}(e_0)}^* \sum_{r \in \mathbf{R}} \iota_r \Gamma_{[r]} \iota_r^* \\ &= L_{\overline{\mathbf{G}}}(Z)X = YL_{\overline{\mathbf{G}}}(Z) = \sum_{s \in \mathbf{S}} \iota_s \Gamma_{[s]} \iota_s^* \iota_{s(e_0)} \Delta_p \iota_{\mathbf{r}(e_0)}^* \\ &= \iota_{s(e_0)} \Gamma_{[s(e_0)]} \iota_{s(e_0)}^* \iota_{s(e_0)} \Delta_p \iota_{\mathbf{r}(e_0)}^* = \iota_{s(e_0)} \Gamma_p \Delta_p \iota_{\mathbf{r}(e_0)}^*. \end{aligned}$$

This proves that $\Delta_p \Gamma_p = \Gamma_p \Delta_p$. Since Δ_p is an arbitrary element of Δ_p and $p \in \mathbf{P}$ was also chosen arbitrarily, we obtain that $\Gamma \in \Delta'_p$ for each $p \in \mathbf{P}$.

One easily verifies that the pair (X, Y) with X and Y as in (3.5) where $\Gamma_p \in \Delta'_p$ for each $p \in \mathbf{P}$ is in $\Delta'_{\overline{\mathbf{G}}}$. Hence the proof is complete. \square

Now suppose that we are given an operator $M \in \mathcal{L}(\mathcal{H}_{\mathbf{S}}, \mathcal{H}_{\mathbf{R}})$ along with the A -graph $\overline{\mathbf{G}} = (G, \{\Delta_p \subset \mathcal{L}(\mathcal{H}_p)\}_{p \in \mathbf{P}})$ as above. We then define the $\mu_{\overline{\mathbf{G}}}$ -structured singular value of M as in (1.2) but with $\Delta_{\overline{\mathbf{G}}}$ as in (3.3) in place of Δ :

$$\mu_{\Delta_{\overline{\mathbf{G}}}}(M) = \frac{1}{\inf\{\|\Delta\| : \Delta \in \Delta_{\overline{\mathbf{G}}}, 1 \in \sigma(M\Delta)\}}. \quad (3.6)$$

The analogue of the D -scaled version of μ is defined as

$$\widehat{\mu}_{\Delta_{\overline{\mathbf{G}}}}(M) = \inf\{\|XMY^{-1}\| : (X, Y) \in \Delta'_{\overline{\mathbf{G}}} \text{ with } X, Y \text{ invertible}\}. \quad (3.7)$$

As in the classical case, $\widehat{\mu}_{\Delta_{\overline{\mathbf{G}}}}(M)$ has the following properties:

- Computation of $\widehat{\mu}_{\Delta_{\overline{\mathbf{G}}}}(M)$ can be reduced to a C^* -algebra LOI (Linear Operator Inequality) computation: $\widehat{\mu}_{\Delta_{\overline{\mathbf{G}}}}(M) < 1$ if and only if there exists a positive definite structured solution $(X, Y) \in \Delta'_{\overline{\mathbf{G}}}$ of the structured Stein inequality

$$M^*XM - Y \prec 0. \quad (3.8)$$

In cases of interest, the C^* -algebra is concretely identified as a subspace of structured finite matrices and the structured LOI becomes a structured LMI (Linear Matrix Inequality).

- $\widehat{\mu}_{\Delta_{\overline{\mathbf{G}}}}(M)$ is always an upper bound for $\mu_{\Delta_{\overline{\mathbf{G}}}}(M)$:

$$\mu_{\Delta_{\overline{\mathbf{G}}}}(M) \leq \widehat{\mu}_{\Delta_{\overline{\mathbf{G}}}}(M). \quad (3.9)$$

Rather than pursuing this general situation further, we now discuss two particular special cases which will be our focus for the rest of the paper.

3.2. The classical uncertainty structure. Let us now suppose that we are given an M-graph $(G, \{\mathcal{H}_p\}_{p \in \mathbf{P}})$ and we take the C^* -subalgebra of $\mathcal{L}(\mathcal{H}_p)$ to be simply $\Delta_p = \{sI_{\mathcal{H}_p} : s \in \mathbb{C}\}$. Then a $Z \in \Delta^{\mathbf{E}}$ has the form $Z = (Z_e)_{e \in \mathbf{E}}$ where $Z_e = \lambda_e I_{\mathcal{H}_p}$ for complex numbers λ_e . Rather than write

$$L_{\overline{\mathbf{G}}}(Z) = \sum_{e \in \mathbf{E}} \iota_{\mathbf{s}(e)}(\lambda_e I_{\mathcal{H}_{[\mathbf{r}(e)]}}) \iota_{\mathbf{r}(e)}^*,$$

we may write $L_{\overline{\mathbf{G}}}(Z)$ directly as a function of the tuple $(\lambda_e)_{e \in \mathbf{E}}$ of complex numbers:

$$L_{\overline{\mathbf{G}}}(Z) = L_{\mathbf{G}}(\lambda) := \sum_{e \in \mathbf{E}} \lambda_e L_{\mathbf{G},e} \text{ where } L_{\mathbf{G},e} = \iota_{\mathbf{s}(e)} \iota_{\mathbf{r}(e)}^* \text{ for } e \in \mathbf{E}. \quad (3.10)$$

Let us write more simply

$$\Delta_{\mathbf{G}} = \{L_{\mathbf{G}}(\lambda) : \lambda = (\lambda_e)_{e \in \mathbf{E}}, \lambda_e \in \mathbb{C}\} \quad (3.11)$$

for the associated uncertainty structure $\Delta_{\overline{\mathbf{G}}}$ with this special choice of C^* -subalgebras $\Delta_p = \{sI_{\mathcal{H}_p} : s \in \mathbb{C}\}$. Note next that in this case $\Delta'_p = \mathcal{L}(\mathcal{H}_p)$. We therefore read off from Proposition 3.1 that the intertwining space $\Delta'_{\mathbf{G}} := \Delta'_{\overline{\mathbf{G}}}$ is given by

$$\Delta'_{\mathbf{G}} := \{(X, Y) : X = \sum_{r \in \mathbf{R}} \iota_r \Gamma_{[r]} \iota_r^*, Y = \sum_{s \in \mathbf{S}} \iota_s \Gamma_{[s]} \iota_s^* \text{ where } \Gamma_p \in \mathcal{L}(\mathcal{H}_p), p \in \mathbf{P}\}. \quad (3.12)$$

To make $\Delta_{\mathbf{G}}$ more explicit, it is convenient to introduce some auxiliary notation. We let $\tilde{\mathcal{H}}_s = \mathbb{C}$ for each source vertex $s \in \mathbf{S}$ and similarly $\tilde{\mathcal{H}}_r = \mathbb{C}$ for each range vertex $r \in \mathbf{R}$. For each connected component $p \in \mathbf{P}$, we let

$$\tilde{\mathcal{H}}_{\mathbf{S},p} = \bigoplus_{s \in \mathbf{S}_p} \tilde{\mathcal{H}}_s, \quad \tilde{\mathcal{H}}_{\mathbf{R},p} = \bigoplus_{r \in \mathbf{R}_p} \tilde{\mathcal{H}}_r$$

and finally

$$\tilde{\mathcal{H}}_{\mathbf{S}} = \bigoplus_{p \in \mathbf{P}} \tilde{\mathcal{H}}_{\mathbf{S},p}, \quad \tilde{\mathcal{H}}_{\mathbf{R}} = \bigoplus_{p \in \mathbf{P}} \tilde{\mathcal{H}}_{\mathbf{R},p}.$$

Note that these spaces amount to the quantities in (3.1) in the case of the multiplicity-one assignment $\mathcal{H}_p = \mathbb{C}$ for each component p of the graph G ; in general we have the tensor factorizations

$$\mathcal{H}_{\mathbf{R},p} = \tilde{\mathcal{H}}_{\mathbf{R},p} \otimes \mathcal{H}_p, \quad \mathcal{H}_{\mathbf{S},p} = \tilde{\mathcal{H}}_{\mathbf{S},p} \otimes \mathcal{H}_p. \quad (3.13)$$

Then it is not difficult to see that the uncertainty structure (3.11) can be written more explicitly as

$$\Delta_{\mathbf{G}} = \left\{ \bigoplus_{p \in \mathbf{P}} W_p \otimes I_{\mathcal{H}_p} : W_p \in \mathcal{L}(\tilde{\mathcal{H}}_{\mathbf{R},p}, \tilde{\mathcal{H}}_{\mathbf{S},p}) \right\}. \quad (3.14)$$

Since G is a finite graph, by assumption, we can number the path-connected components p_1, \dots, p_K , with $K = \#(\mathbf{P}) < \infty$. When convenient we shall use k as an index rather than p_k when referring to elements associated with the k -th connected component. Say the k -th connected component p_k has n_k source vertices and m_k range vertices. We then number the source vertices $s_{k,i}$ and range vertices $r_{k,j}$ for

$i = 1 \dots, n_k$ and $j = 1 \dots, m_k$ and write $e_{k,ij}$ for the edge connecting source vertex $s_{k,i}$ to range vertex $r_{k,j}$. Thus we have the following labelings:

$$\begin{aligned}\mathbf{S} &= \cup_{k=1}^K \mathbf{S}_k \text{ where } \mathbf{S}_k = \{s_{k,i} : 1 \leq i \leq n_k\}, \\ \mathbf{R} &= \cup_{k=1}^K \mathbf{R}_k \text{ where } \mathbf{R}_k = \{r_{k,j} : 1 \leq j \leq m_k\}, \\ \mathbf{E} &= \cup_{k=1}^K \mathbf{E}_k \text{ where } \mathbf{E}_k = \{e_{k,ij} : 1 \leq i \leq n_k, 1 \leq j \leq m_k\}.\end{aligned}$$

Then the uncertainty structure (3.11) now assumes the form

$$\Delta_{\mathbf{G}} = \left\{ \sum_{k,i,j} \lambda_{k,i,j} \iota_{s_{k,i}} \iota_{r_{k,j}}^* : \lambda_{k,i,j} \in \mathbb{C} \text{ arbitrary} \right\}$$

with the more explicit formulation (3.14) becoming

$$\Delta_{\mathbf{G}} = \{ \text{diag}_{k=1,\dots,K} W_k \otimes I_{\mathcal{H}_k} : W_k \in \mathbb{C}^{n_k \times m_k} \}. \quad (3.15)$$

In case all \mathcal{H}_k are finite dimensional, tensoring with $I_{\mathcal{H}_k}$ just says that each Δ_k is allowed to have multiplicity equal to $\dim \mathcal{H}_k$. We note that the structure (1.1) discussed in Section 1 is the special case where $n_k = m_k$ for all k and $\dim \mathcal{H}_k = 1$ whenever $n_k = m_k > 1$.

3.3. The enhanced classical uncertainty structure. We now describe a second special form for an A-graph. Suppose that we are given an M-graph $(G, \{\mathcal{H}_p : p \in \mathbf{P}\})$ where \mathcal{H}_p has the tensor-product form $\mathcal{H}_p = \mathcal{K} \otimes \mathcal{H}_p^\circ$ for a fixed Hilbert space \mathcal{K} and coefficient Hilbert spaces \mathcal{H}_p° . It will be convenient to have a notation also for the M-graph with coefficient Hilbert spaces \mathcal{H}_p° :

$$\mathbf{G}^\circ = (G, \{\mathcal{H}_p^\circ : p \in \mathbf{P}\}).$$

We now specify the C^* -subalgebra $\Delta_p \subset \mathcal{L}(\mathcal{H}_p)$ to be

$$\Delta_p = \mathcal{L}(\mathcal{K}) \otimes I_{\mathcal{H}_p^\circ},$$

and denote the associated A-graph by $\overline{\mathbf{G}}$. If $Z' = (Z'_e)_{e \in \mathbf{E}}$ is an element of $\Delta^{\mathbf{E}}$, then each Z'_e has the form

$$Z'_e = Z_e \otimes I_{\mathcal{H}_p^\circ}$$

where Z_e is an arbitrary operator on \mathcal{K} . Then the operator

$$L_{\overline{\mathbf{G}}}(Z') = \sum_{e \in \mathbf{E}} \iota_{\mathbf{s}(e)} (Z_e \otimes I_{\mathcal{H}_p^\circ}) \iota_{\mathbf{r}(e)}^*$$

is really a function $L_{\mathbf{G}}(Z)$ of the \mathbf{E} -tuple $Z = (Z_e)_{e \in \mathbf{E}}$ of operators on \mathcal{K} . If we let $L_{\mathbf{G}^\circ}(z)$ be as in Subsection 3.2 associated with the M-graph \mathbf{G}° , with the \circ -super index carried over in the notation,

$$L_{\mathbf{G}^\circ}(\lambda) = \sum_{e \in \mathbf{E}} \lambda_e L_{\mathbf{G}^\circ, e} \quad \text{where} \quad L_{\mathbf{G}^\circ, e} = \iota_{\mathbf{s}(e)}^\circ (\iota_{\mathbf{r}(e)}^\circ)^*,$$

then, by using the identities

$$\iota_{\mathbf{s}(e)} = I_{\mathcal{K}} \otimes \iota_{\mathbf{s}(e)}^\circ, \quad \iota_{\mathbf{r}(e)} = I_{\mathcal{K}} \otimes \iota_{\mathbf{r}(e)}^\circ,$$

it is easily verified that

$$L_{\overline{\mathbf{G}}}(Z') = L_{\mathbf{G}^\circ}(Z) := \sum_{e \in \mathbf{E}} Z_e \otimes L_{\mathbf{G}^\circ, e}. \quad (3.16)$$

More explicitly, in the notation used at the end of Subsection 3.2, we see that we have the enhanced versions of the factorizations (3.13)

$$\mathcal{H}_{\mathbf{R},p} = \mathcal{K} \otimes \mathcal{H}_{\mathbf{R},p}^\circ = \mathcal{K} \otimes \tilde{\mathcal{H}}_{\mathbf{R},p} \otimes \mathcal{H}_p^\circ, \quad \mathcal{H}_{\mathbf{S},p} = \mathcal{K} \otimes \mathcal{H}_{\mathbf{S},p}^\circ = \mathcal{K} \otimes \tilde{\mathcal{H}}_{\mathbf{S},p} \otimes \mathcal{H}_p^\circ \quad (3.17)$$

and the associated uncertainty structure $\Delta_{\overline{\mathbf{G}}}$ can be presented as follows:

$$\Delta_{\overline{\mathbf{G}}} = \left\{ \bigoplus_{p \in \mathbf{P}} W_p \otimes I_{\mathcal{H}_p^\circ} : W_p \in \mathcal{L}(\mathcal{K} \otimes \tilde{\mathcal{H}}_{\mathbf{R},p}, \mathcal{K} \otimes \tilde{\mathcal{H}}_{\mathbf{S},p}) \right\} \quad (3.18)$$

or in matrix form,

$$\Delta_{\overline{\mathbf{G}}} = \{W = \text{diag}_{k=1,\dots,K} [W_k \otimes I_{\mathcal{H}_{p_k}^\circ}] : W_k \in \mathcal{L}(\mathcal{K})^{n_k \times m_k}\}. \quad (3.19)$$

We shall be interested in computing $\mu_{\Delta_{\overline{\mathbf{G}}}}(M)$ for the case where M has the tensored form $M = I_{\mathcal{K}} \otimes M^\circ$ for an operator $M^\circ \in \mathcal{L}(\mathcal{H}_{\mathbf{S}}^\circ, \mathcal{H}_{\mathbf{R}}^\circ)$. It is then natural to use the shorthand notation

$$\tilde{\mu}_{\mathbf{G}^\circ}(M^\circ) := \mu_{\Delta_{\overline{\mathbf{G}}}}(I_{\mathcal{K}} \otimes M^\circ).$$

For $\Delta_p = \mathcal{L}(\mathcal{K}) \otimes I_{\mathcal{H}_p^\circ}$, we have

$$\Delta'_p = I_{\mathcal{K}} \otimes \mathcal{L}(\mathcal{H}_p^\circ).$$

and hence we read off from Proposition 3.1 that

$$\begin{aligned} \Delta'_{\overline{\mathbf{G}}} &= \{(X, Y) : X = \sum_{r \in \mathbf{R}} I_{\mathcal{K}} \otimes \iota_r \Gamma_{[r]}^\circ \iota_r^*, Y = \sum_{s \in \mathbf{S}} I_{\mathcal{K}} \otimes \iota_s \Gamma_{[s]}^\circ \iota_s^* \text{ where } \Gamma_p^\circ \in \mathcal{L}(\mathcal{H}_p^\circ)\} \\ &= I_{\mathcal{K}} \otimes \Delta'_{\mathbf{G}^\circ}. \end{aligned} \quad (3.20)$$

3.4. Main Result. We can now state our Main Result as follows.

Theorem 3.2 (Main Result). *Let $\overline{\mathbf{G}}$ and \mathbf{G}° be as in Subsection 3.3 with \mathcal{K} taken to be an infinite-dimensional separable Hilbert and all \mathcal{H}_p° finite dimensional, where $p \in \mathbf{P}$. Then, for any linear operator*

$$M^\circ : \bigoplus_{p \in \mathbf{P}} \mathcal{H}_{\mathbf{S},p}^\circ = \bigoplus_{p \in \mathbf{P}} (\tilde{\mathcal{H}}_{\mathbf{S},p} \otimes \mathcal{H}_p^\circ) \rightarrow \bigoplus_{p \in \mathbf{P}} \mathcal{H}_{\mathbf{R},p}^\circ = \bigoplus_{p \in \mathbf{P}} (\tilde{\mathcal{H}}_{\mathbf{R},p} \otimes \mathcal{H}_p^\circ)$$

we have

$$\tilde{\mu}_{\Delta_{\mathbf{G}^\circ}}(M^\circ) := \mu_{\Delta_{\overline{\mathbf{G}}}}(I_{\mathcal{K}} \otimes M^\circ) = \hat{\mu}_{\Delta_{\mathbf{G}^\circ}}(M^\circ).$$

In particular

$$\tilde{\mu}_{\Delta_{\mathbf{G}^\circ}}(M^\circ) < 1 \iff \hat{\mu}_{\Delta_{\mathbf{G}^\circ}}(M^\circ) < 1$$

and testing whether $\tilde{\mu}_{\Delta_{\mathbf{G}^\circ}}(M^\circ) < 1$ reduces to a finite-dimensional LMI.

As explained in the Introduction, in the succeeding sections we discuss two distinct approaches to this result: one based on the earlier work of Ball-Groenewald-Malakorn [11], the other on the work of Dullerud-Paganini [35, 22].

We conclude this section with a remark that reduces the claims of Theorem 3.2 to a single implication.

Remark 3.3. We first observe that the inequality $\tilde{\mu}_{\Delta_{\mathbf{G}^\circ}}(M^\circ) \leq \hat{\mu}_{\Delta_{\mathbf{G}^\circ}}(M^\circ)$ holds. This follows from two observations. Firstly, we have the inequality $\tilde{\mu}_{\Delta_{\mathbf{G}^\circ}}(M^\circ) = \mu_{\Delta_{\overline{\mathbf{G}}}}(I_{\mathcal{K}} \otimes M^\circ) \leq \hat{\mu}_{\Delta_{\overline{\mathbf{G}}}}(I_{\mathcal{K}} \otimes M^\circ)$, as observed on the level of Subsection 3.1 on Page 9. Secondly, since $\Delta'_{\overline{\mathbf{G}}} = I_{\mathcal{K}} \otimes \Delta'_{\mathbf{G}^\circ}$, by (3.20), we have

$$\|X(I_{\mathcal{K}} \otimes M^\circ)Y^{-1}\| = \|X^\circ M^\circ (Y^\circ)^{-1}\|$$

for any $(X, Y) = (I_{\mathcal{K}} \otimes X^\circ, I_{\mathcal{K}} \otimes Y^\circ) \in \Delta'_{\mathbf{G}}$ with $(X^\circ, Y^\circ) \in \Delta'_{\mathbf{G}^\circ}$. Consequently, we obtain $\hat{\mu}_{\Delta_{\mathbf{G}}} (I_{\mathcal{K}} \otimes M^\circ) = \hat{\mu}_{\Delta_{\mathbf{G}^\circ}} (M^\circ)$, which yields the claimed inequality. Hence it remains to prove $\hat{\mu}_{\Delta_{\mathbf{G}^\circ}} (M^\circ) \leq \tilde{\mu}_{\Delta_{\mathbf{G}^\circ}} (M^\circ)$. By a scaling argument, this in turn reduces to showing:

$$\tilde{\mu}_{\mathbf{G}^\circ} (M^\circ) < 1 \implies \hat{\mu}_{\mathbf{G}^\circ} (M^\circ) < 1. \quad (3.21)$$

4. NONCOMMUTATIVE BOUNDED REAL LEMMA, STATE-SPACE SIMILARITY THEOREM, AND STRUCTURED SINGULAR VALUE VERSUS DIAGONAL SCALING

Throughout this section, let \mathbf{G} be a M-graph:

$$\mathbf{G} = (G, \{\mathcal{H}_p : p \in \mathbf{P}\}).$$

Here we give a proof of our Main Result (Theorem 3.2) based on two theorems from [10, 11] regarding the Schur-Agler class and colligation matrices associated with the M-graph \mathbf{G} .

For this purpose we let $z = (z_e)_{e \in \mathbf{E}}$ be a collection of freely noncommuting indeterminates indexed by the edge set \mathbf{E} . We let $L_{\mathbf{G}}(z)$ be the formal linear pencil

$$L_{\mathbf{G}}(z) := \sum_{e \in \mathbf{E}} z_e L_{\mathbf{G},e} \quad (4.1)$$

where the coefficients $L_{\mathbf{G},e}$ are as in (3.10). For $Z = (Z_e)_{e \in \mathbf{E}}$ a tuple of operators on some auxiliary Hilbert space \mathcal{K} , we evaluate the formal pencil $L_{\mathbf{G}}(z)$ at the argument Z by using tensor products just as in (3.16):

$$L_{\mathbf{G}}(Z) = \sum_{e \in \mathbf{E}} Z_e \otimes L_{\mathbf{G},e}. \quad (4.2)$$

This framework includes as a special case the situation where $\mathcal{K} = \mathbb{C}$ and each Z_e is an operator on the one-dimensional space \mathbb{C} ; for this case we write $\lambda = (\lambda_e)_{e \in \mathbf{E}}$ with $\lambda_e \in \mathbb{C}$ instead of $Z = (Z_e)_{e \in \mathbf{E}}$ and we arrive at the classical operator pencil in the \mathbf{E} -tuple of complex numbers $\lambda = (\lambda_e)_{e \in \mathbf{E}}$ as in (3.10):

$$L_{\mathbf{G}}(\lambda) = \sum_{e \in \mathbf{E}} \lambda_e L_{\mathbf{G},e}.$$

Before turning to the results from [10, 11] and the proof of Theorem 3.2, we recall some facts about formal power series.

4.1. Formal power series. We let $\mathcal{F}_{\mathbf{E}}$ be the free monoid on the generating set \mathbf{E} , i.e., the free semigroup with the empty word \emptyset serving as the identity element. Thus a generic element α of $\mathcal{F}_{\mathbf{E}}$ has the form $\alpha = e_{i_N} \cdots e_{i_1}$ where $e_{i_j} \in \mathbf{E}$ for each $j = 1, \dots, N$. When $\alpha \in \mathcal{F}_{\mathbf{E}}$ has this form, we say that the *length* $|\alpha|$ of α is N ; we include the empty word \emptyset as an element of $\mathcal{F}_{\mathbf{E}}$, considered to have length zero. Multiplication of two elements $\alpha = e_{i_N} \cdots e_{i_1}$ and $\beta = e_{j_M} \cdots e_{j_1}$ of $\mathcal{F}_{\mathbf{E}}$ is by concatenation:

$$\alpha \cdot \beta = e_{i_N} \cdots e_{i_1} e_{\beta_{j_M}} \cdots e_{\beta_{j_1}}$$

with the empty word \emptyset serving as the identity element of $\mathcal{F}_{\mathbf{E}}$. Furthermore, the transpose α^\top of $\alpha = e_{i_N} \cdots e_{i_1}$ is defined as $\alpha^\top = e_{j_1} \cdots e_{j_M}$. Given the \mathbf{E} -tuple $z = (z_e)_{e \in \mathbf{E}}$ of freely noncommuting indeterminates and an element $\alpha = e_{i_N} \cdots e_{i_1}$ we define the noncommutative monomial z^α by

$$z^\alpha = z_{e_{i_N}} \cdots z_{e_{i_1}}$$

with an individual indeterminate z_e identified with z^α if $\alpha = e$ is a word of length one and with z^\emptyset identified with 1.

For \mathcal{X} a linear space, we let $\mathcal{X}\langle\langle z \rangle\rangle$ denote the set of all formal power series $\sum_{\alpha \in \mathcal{F}_{\mathbf{E}}} x_\alpha z^\alpha$ with coefficients x_α coming from \mathcal{X} . Two formal power series $\sum_{\alpha \in \mathcal{F}_{\mathbf{E}}} x_\alpha z^\alpha$ and $\sum_{\alpha \in \mathcal{F}_{\mathbf{E}}} y_\alpha z^\alpha$ are said to be equal if $x_\alpha = y_\alpha$ for all $\alpha \in \mathbf{E}$. If \mathcal{X}' and \mathcal{X}'' are also linear spaces for which a multiplication $\mathcal{X}' \times \mathcal{X}'' \rightarrow \mathcal{X}''$ is defined and if we are given two formal series $x(z) = \sum_{\alpha \in \mathcal{F}_{\mathbf{E}}} x_\alpha z^\alpha \in \mathcal{X}\langle\langle z \rangle\rangle$ and $x'(z) = \sum_{\beta \in \mathcal{F}_{\mathbf{E}}} x'_\beta z^\beta \in \mathcal{X}'\langle\langle z \rangle\rangle$, then the product formal series $x'(z) \cdot x(z) \in \mathcal{X}''\langle\langle z \rangle\rangle$ is always well defined and given by

$$(x' \cdot x)(z) = \sum_{\gamma \in \mathcal{F}_{\mathbf{E}}} \left(\sum_{\beta, \alpha \in \mathcal{F}_{\mathbf{E}}: \beta \cdot \alpha = \gamma} x'_\beta x_\alpha \right) z^\gamma.$$

Assume \mathcal{X} is endowed with some appropriate topology (typically \mathcal{X} will be a Hilbert space or the space of bounded linear operators between two Hilbert spaces). As is now common in the theory of noncommutative functions (see e.g. [26, 30]), we will often view a formal power series $x(z) = \sum_{\alpha \in \mathcal{F}_{\mathbf{E}}} x_\alpha z^\alpha \in \mathcal{X}\langle\langle z \rangle\rangle$ as a function whose variables are operators on some auxiliary separable Hilbert space \mathcal{K} . In this way, for an \mathbf{E} -tuple $Z = (Z_e)_{e \in \mathbf{E}}$ of linear operators acting on \mathcal{K} and a formal power series $x(z) = \sum_{\alpha \in \mathcal{F}_{\mathbf{E}}} x_\alpha z^\alpha$ we define an element $x(Z) \in \mathcal{L}(\mathcal{K}) \otimes \mathcal{X}$ by

$$x(Z) = \sum_{\alpha \in \mathcal{F}_{\mathbf{E}}} Z^\alpha \otimes x_\alpha \in \mathcal{L}(\mathcal{K}) \otimes \mathcal{X} \quad (4.3)$$

whenever the series converges in the appropriate topology of $\mathcal{L}(\mathcal{K}) \otimes \mathcal{X}$. Here we use the notation

$$Z^\alpha = Z_{e_{i_N}} \cdots Z_{e_{i_1}} \in \mathcal{L}(\mathcal{K}) \quad \text{for } \alpha = e_{i_N} \cdots e_{i_1} \in \mathcal{F}_{\mathbf{E}}.$$

Notice that the point evaluation in (4.3) generalizes the one already introduced for the linear case in (3.16).

4.2. The Schur-Agler class and colligation matrices associated with \mathbf{G} .

Let \mathcal{U} and \mathcal{Y} be two auxiliary Hilbert spaces. Given a formal power series $S(z) = \sum_{\alpha \in \mathcal{F}_{\mathbf{E}}} S_\alpha z^\alpha \in \mathcal{L}(\mathcal{U}, \mathcal{Y})\langle\langle z \rangle\rangle$, we say that S is in the Schur-Agler class $\mathcal{SA}_{\mathbf{G}}(\mathcal{U}, \mathcal{Y})$ associated with the M-graph \mathbf{G} if for any \mathbf{E} -tuple $Z = (Z_e)_{e \in \mathbf{E}}$ of operators $Z_e \in \mathcal{L}(\mathcal{K})$ such that $\|L_{\mathbf{G}}(Z)\| < 1$, the evaluation $S(Z)$ via (4.3) is in $\mathcal{L}(\mathcal{K} \otimes \mathcal{U}, \mathcal{K} \otimes \mathcal{Y})$ and satisfies $\|S(Z)\| \leq 1$. We note that the test-class of \mathbf{E} -tuples $Z = (Z_e)_{e \in \mathbf{E}}$ is independent of the choice of multiplicity structure for \mathbf{G} , as changing the multiplicity structure of \mathbf{G} does not effect the norm $\|L_{\mathbf{G}}(Z)\|$. For purposes of defining the Schur-Agler class, we may as well assume that the underlying graph G is taken with multiplicity-1 structure ($\mathcal{H}_p = \mathbb{C}$ for each p), and we write $\mathcal{SA}_G(\mathcal{U}, \mathcal{Y})$ rather than $\mathcal{SA}_{\mathbf{G}}(\mathcal{U}, \mathcal{Y})$.

The following result was obtained in [10]

Theorem 4.1. (See [10, Theorem 5.3].) *Give a formal power series $S \in \mathcal{L}(\mathcal{U}, \mathcal{Y})\langle\langle z \rangle\rangle$, $S(z) = \sum_{\alpha \in \mathcal{F}_{\mathbf{E}}} S_\alpha z^\alpha$, the following conditions are equivalent:*

- (1) *S is in the Schur-Agler class $\mathcal{SA}_G(\mathcal{U}, \mathcal{Y})$.*
- (2) *There is a multiplicity assignment $\{\mathcal{H}_p: p \in \mathbf{P}\}$ giving rise to an M-graph $\mathbf{G} = (G, \{\mathcal{H}_p: p \in \mathbf{P}\})$ and a formal power series $H \in \mathcal{L}(\mathcal{H}_{\mathbf{S}}, \mathcal{Y})\langle\langle z \rangle\rangle$ so*

that S has the Agler decomposition

$$I - S(z)S(w)^* = H(z)(I - L_{\mathbf{G}}(z)L_{\mathbf{G}}(w)^*)H(w)^*. \quad (4.4)$$

Here $\bar{w} = (\bar{w}_e)_{e \in \mathbf{E}}$ is another \mathbf{E} -tuple of freely noncommuting indeterminates, we set $H(w)^* = \sum_{\beta \in \mathcal{F}_{\mathbf{E}}} (H_{\beta})^* \bar{w}^{\beta^\top}$ if $H(z) = \sum_{\alpha \in \mathcal{F}_{\mathbf{E}}} H_{\alpha} z^{\alpha}$ and define $S(w)^*$ accordingly, and (4.4) is to be interpreted as an formal power series in the $\mathbf{E} \dot{\cup} \mathbf{E}$ -tuple $(z_e)_{e \in \mathbf{E}} \cup (\bar{w}_e)_{e \in \mathbf{E}}$.

- (3) S has a dissipative noncommutative structured realization, i.e., there exists a multiplicity assignment $\{\mathcal{H}_p : p \in \mathbf{P}\}$ with associated M -graph

$$\mathbf{G} = (G, \{\mathcal{H}_p : p \in \mathbf{P}\})$$

together with a contractive colligation matrix

$$\mathbf{U} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} : \begin{bmatrix} \mathcal{H}_{\mathbf{S}} \\ \mathcal{U} \end{bmatrix} \rightarrow \begin{bmatrix} \mathcal{H}_{\mathbf{R}} \\ \mathcal{Y} \end{bmatrix} \quad (4.5)$$

so that

$$S(z) = D + C(I - L_{\mathbf{G}}(z)A)^{-1}L_{\mathbf{G}}(z)B. \quad (4.6)$$

If we are given a colligation matrix \mathbf{U} as in (4.5) and define the associated formal power series $S(z)$ via (4.6), then it is possible that S is in the Schur-Agler class even though the colligation matrix \mathbf{U} is not contractive; indeed, a sufficient condition which is weaker than contractivity of \mathbf{U} is that there exist an invertible change-of-basis matrix Γ_p on \mathcal{H}_p for each connected component $p \in \mathbf{P}$ of \mathbf{G} so that the transformed colligation matrix

$$\mathbf{U}' = \begin{bmatrix} A' & B' \\ C' & D' \end{bmatrix} := \begin{bmatrix} \bigoplus_{r \in \mathbf{R}} \Gamma_{[r]} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} \bigoplus_{s \in \mathbf{S}} (\Gamma_{[s]})^{-1} & 0 \\ 0 & I \end{bmatrix}$$

is a contraction:

$$\left\| \begin{bmatrix} \bigoplus_{r \in \mathbf{R}} \Gamma_{[r]} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} \bigoplus_{s \in \mathbf{S}} (\Gamma_{[s]})^{-1} & 0 \\ 0 & I \end{bmatrix} \right\| \leq 1. \quad (4.7)$$

Equivalently, one can ask for positive definite matrices $\Gamma_p \succ 0$ on each partial state space \mathcal{H}_p so that

$$\begin{bmatrix} A^* & C^* \\ B^* & D^* \end{bmatrix} \begin{bmatrix} \bigoplus_{r \in \mathbf{R}} \Gamma_{[r]} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} - \begin{bmatrix} \bigoplus_{s \in \mathbf{S}} \Gamma_{[s]} & 0 \\ 0 & I \end{bmatrix} \preceq 0. \quad (4.8)$$

If we assume that all the spaces \mathcal{H}_p are finite-dimensional and also impose a structured minimality assumption, this sufficient condition is also necessary (see Theorem 3.1 in [11]). A result of this type is known as a *Bounded Real Lemma* (see e.g. [45]). The idea of a *strict Bounded Real Lemma* (see e.g. [37] and Lemma 7.4 in [22]) is to trade in the minimality assumption for a stability assumption. The Bounded Real Lemma in the context of SNMLSs is the following result.

Theorem 4.2. (See [11, Theorem 3.4].) *Suppose that we are given an A -graph of the form $\mathbf{G} = (G, \{\Delta_p = \{sI_{\mathcal{H}_p} : s \in \mathbb{C}\} \subset \mathcal{L}(\mathcal{H}_p)\})$, where \mathcal{H}_p is a finite-dimensional Hilbert space for each $p \in \mathbf{P}$, together with a colligation matrix \mathbf{U} as in (4.5). Associate with \mathbf{U} the formal power series $S(z)$ as in (4.6). Then the following conditions are equivalent:*

(1) (i) A is uniformly \mathbf{G} -stable:

$$\sup_{Z: \|L_{\mathbf{G}}(Z)\| \leq 1} \|(I - L_{\mathbf{G}}(Z)A)^{-1}\| < \infty$$

and (ii) there exists a $\rho < 1$ so that $S \in \rho \cdot \mathcal{SA}_G(\mathcal{U}, \mathcal{Y})$:

$$\sup_{Z: \|L_{\mathbf{G}}(Z)\| \leq 1} \|S(Z)\| \leq \rho < 1.$$

(2) There exist invertible matrices Γ_p on \mathcal{H}_p , for each $p \in \mathbf{P}$, so that the strict version of condition (4.7) holds:

$$\left\| \begin{bmatrix} \bigoplus_{r \in \mathbf{R}} \Gamma_{[r]} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} \bigoplus_{s \in \mathbf{S}} (\Gamma_{[s]})^{-1} & 0 \\ 0 & I \end{bmatrix} \right\| < 1. \quad (4.9)$$

(3) There exist strictly positive definite operators Γ_p on \mathcal{H}_p , for each $p \in \mathbf{P}$, so that the strict version of (4.8) holds:

$$\begin{bmatrix} A^* & C^* \\ B^* & D^* \end{bmatrix} \begin{bmatrix} \bigoplus_{r \in \mathbf{R}} \Gamma_{[r]} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} - \begin{bmatrix} \bigoplus_{s \in \mathbf{S}} \Gamma_{[s]} & 0 \\ 0 & I \end{bmatrix} \prec 0. \quad (4.10)$$

4.3. Proof of Theorem 3.2. For the remainder of this section we follow the notation of Subsections 3.3 and 3.4. Hence, we consider an M-graph $(G, \{\mathcal{H}_p: p \in \mathbf{P}\})$ where each Hilbert space \mathcal{H}_p has the tensored form $\mathcal{H}_p = \mathcal{K} \oplus \mathcal{H}_p^\circ$ with \mathcal{K} and \mathcal{H}_p° Hilbert spaces, \mathcal{K} separable and \mathcal{H}_p° finite dimensional. As in Subsection 3.3, with this M-graph we associate the M-graph $\mathbf{G}^\circ = (G, \{\mathcal{H}_p^\circ: p \in \mathbf{P}\})$ and the A-graph

$$\overline{\mathbf{G}} = (G, \{\Delta_p = \mathcal{L}(\mathcal{K}) \otimes I_{\mathcal{H}_p^\circ} \subset \mathcal{L}(\mathcal{H}_p), p \in \mathbf{P}\}).$$

The linear pencils $L_{\mathbf{G}}(\lambda)$ and $L_{\mathbf{G}}(Z)$ from Subsection 4.2 then coincide with $L_{\mathbf{G}^\circ}(\lambda)$ and $L_{\mathbf{G}^\circ}(Z) = L_{\overline{\mathbf{G}}}(\lambda)$, respectively, as defined in Subsection 3.3. We proceed here with the notation of Subsection 3.3, i.e., with $L_{\mathbf{G}^\circ}(\lambda)$ and $L_{\mathbf{G}^\circ}(Z)$, as well as the formal pencil $L_{\mathbf{G}^\circ}(z)$ as in (4.2).

Now let us suppose we are given a matrix $M^\circ \in \mathcal{L}(\mathcal{H}_{\mathbf{S}}^\circ, \mathcal{H}_{\mathbf{R}}^\circ)$, where

$$\begin{aligned} \mathcal{H}_{\mathbf{S}}^\circ &= \bigoplus_{p \in \mathbf{P}} \tilde{\mathcal{H}}_{\mathbf{S},p} \otimes \mathcal{H}_p^\circ = \bigoplus_{p \in \mathbf{P}} (\bigoplus_{s \in \mathbf{S}} \tilde{\mathcal{H}}_s \otimes \mathcal{H}_p^\circ), \\ \mathcal{H}_{\mathbf{R}}^\circ &= \bigoplus_{p \in \mathbf{P}} \tilde{\mathcal{H}}_{\mathbf{R},p} \otimes \mathcal{H}_p^\circ = \bigoplus_{p \in \mathbf{P}} (\bigoplus_{r \in \mathbf{R}} \tilde{\mathcal{H}}_r \otimes \mathcal{H}_p^\circ), \end{aligned}$$

where $\tilde{\mathcal{H}}_s = \tilde{\mathcal{H}}_r = \mathbb{C}$ for each $s \in \mathbf{S}, r \in \mathbf{R}$. As before we set $M = I_{\mathcal{K}} \otimes M^\circ \in \mathcal{L}(\mathcal{K} \otimes \mathcal{H}_{\mathbf{R}}^\circ, \mathcal{K} \otimes \mathcal{H}_{\mathbf{S}}^\circ)$.

For the discussion to follow let us introduce the notation

$$\overline{\mathbf{B}}\Delta_{\mathbf{G}^\circ} = \{L_{\mathbf{G}^\circ}(Z): Z = (Z_e)_{e \in \mathbf{E}}, Z_e \in \mathcal{L}(\mathcal{K}) \text{ with } \|L_{\mathbf{G}^\circ}(Z)\| \leq 1\}.$$

As observed in Remark 3.3, it remains to prove the implication:

$$\tilde{\mu}_{\mathbf{G}^\circ}(M^\circ) < 1 \implies \hat{\mu}_{\mathbf{G}^\circ}(M^\circ) < 1. \quad (4.11)$$

The assumption $\tilde{\mu}_{\mathbf{G}^\circ}(M^\circ) < 1$ implies in particular that

$$(I - L_{\mathbf{G}^\circ}(Z)M) \text{ exists for all } Z \text{ with } \|L_{\mathbf{G}^\circ}(Z)\| \leq 1. \quad (4.12)$$

We note that the formal structured resolvent $(I - L_{\mathbf{G}^\circ}(z)M^\circ)^{-1}$ can be written in realization form (4.6)

$$(I - L_{\mathbf{G}^\circ}(z)M^\circ)^{-1} = I + I \cdot (I - L_{\mathbf{G}^\circ}(z)M^\circ)^{-1} L_{\mathbf{G}^\circ}(z) \cdot M^\circ, \quad (4.13)$$

i.e., in the form (4.6) with $\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} M^\circ & M^\circ \\ I & I \end{bmatrix}$. If condition (4.12) can be strengthened to

$$\sup_{Z \in \mathcal{L}(\mathcal{K})^{\mathbf{E}}: \|L_{\mathbf{G}^\circ}(Z)\| \leq 1} \|(I - L_{\mathbf{G}^\circ}(Z)M)^\circ\|^{-1} < \infty \quad (4.14)$$

then condition (i) in statement (1) of Theorem 4.2 (with \mathbf{G} replaced by \mathbf{G}°) is satisfied with M° in place of A . Moreover, if (4.14) holds and if we chose a positive number r slightly larger than the supremum in (4.14), then the power series $S(z) = \frac{1}{r} \cdot (I - L_{\mathbf{G}^\circ}(z)M^\circ)^{-1}$ meets condition (ii) in statement (1) of Theorem 4.2. From the formula (4.13) we see that this $S(z)$ has a realization (4.6) with

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} M^\circ & M^\circ \\ \frac{1}{r}I & \frac{1}{r}I \end{bmatrix}.$$

We may then use the implication (1) \Rightarrow (3) in Theorem 4.2 to conclude that there exist strictly positive definite $\Gamma_p \succ 0$ on \mathcal{H}_p ($p \in \mathbf{P}$) so that

$$\begin{bmatrix} (M^\circ)^* & \frac{1}{r}I \\ (M^\circ)^* & \frac{1}{r}I \end{bmatrix} \begin{bmatrix} \bigoplus_{r \in \mathbf{R}} \Gamma_{[r]} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} M^\circ & M^\circ \\ \frac{1}{r}I & \frac{1}{r}I \end{bmatrix} - \begin{bmatrix} \bigoplus_{s \in \mathbf{S}} \Gamma_{[s]} & 0 \\ 0 & I \end{bmatrix} \prec 0.$$

In particular, peeling off the (1, 1)-entry in this block-matrix inequality yields

$$(M^\circ)^* \left(\bigoplus_{r \in \mathbf{R}} \Gamma_{[r]} \right) M^\circ - \bigoplus_{s \in \mathbf{S}} \Gamma_{[s]} \prec 0$$

from which we read off that $\widehat{\mu}_{\mathbf{G}^\circ}(M^\circ) < 1$ as required. This analysis completes a proof of Theorem 3.2 pending a justification for the jump from (4.12) to (4.14).

We note that without loss of generality we may take the separable infinite-dimensional Hilbert space \mathcal{K} to be ℓ^2 (the space of square-summable complex-valued sequences indexed by the nonnegative integers \mathbb{Z}_+). We conclude that the following lemma, when specialized to the case $M = I_{\ell^2} \otimes M^\circ$ and combined with the analysis in the previous discussion, leads to a complete proof of Theorem 3.2. The construction of the key operator \widehat{W} in the proof adapts ideas from the proof of Proposition B.1 in [22] which can be traced further back to the work of Shamma [41].

Lemma 4.3. *Let $M \in \mathcal{L}(\ell^2 \otimes \mathcal{H}_{\mathbf{S}}^\circ, \ell^2 \otimes \mathcal{H}_{\mathbf{R}}^\circ)$ be shift invariant: $MV_{\mathbf{S}} = V_{\mathbf{R}}M$ where we set $V_{\mathbf{R}} = V \otimes I_{\mathcal{H}_{\mathbf{R}}^\circ}$, $V_{\mathbf{S}} = V \otimes I_{\mathcal{H}_{\mathbf{S}}^\circ}$ where V is the unilateral shift operator on ℓ^2 :*

$$V: (x_0, x_1, \dots) \mapsto (0, x_0, x_1, \dots).$$

Assume that the inverse $(I - L_{\mathbf{G}^\circ \text{irc}}(Z)M)^\circ$ exists for all \mathbf{E} -tuples $Z = (Z_e)_{e \in \mathbf{E}}$ in $\mathcal{L}(\ell^2)$ such that $\|L_{\mathbf{G}^\circ}(Z)\| \leq 1$. Then the collection of all such inverses is uniformly bounded:

$$\sup\{\|(I - L_{\mathbf{G}^\circ}(Z)M)^\circ\|: Z = (Z_e)_{e \in \mathbf{E}}, Z_e \in \mathcal{L}(\ell^2) \text{ with } \|L_{\mathbf{G}^\circ}(Z)\| \leq 1\} < \infty. \quad (4.15)$$

Proof. For integers $0 \leq n_0 \leq N$, let $\ell^2[n_0, N]$ denote the subspace of sequences in ℓ^2 with support in the positions indexed by n_0, \dots, N ; similarly $\ell^2[n_0, N)$ and $\ell^2[n_0, \infty)$ stand for the subspaces ℓ^2 with support in $n_0, \dots, N-1$ and n_0, \dots . As a matter of notation we write $P_{[n_0, N]}$ for the orthogonal projection of $\ell^2 \otimes \mathcal{X}$ onto $\ell^2[n_0, N] \otimes \mathcal{X}$ (where \mathcal{X} is either $\mathcal{H}_{\mathbf{S}}^\circ$ or $\mathcal{H}_{\mathbf{R}}^\circ$ depending on the context); when $n_0 = 0$ we write more simply P_N rather than $P_{[0, N]}$.

We proceed by contradiction. Suppose that $(I - \Delta M)^{-1}$ exists for all $\Delta \in \overline{\mathcal{B}}\Delta_{\mathbf{G}^\circ}$ but that the supremum in (4.15) is infinite. Fix any sequence of positive numbers $\epsilon_n > 0$ such that $\lim_{n \rightarrow \infty} \epsilon_n = 0$. Then we can find $\Delta^{(n)} \in \overline{\mathcal{B}}\Delta_{\mathbf{G}^\circ}$, i.e.,

$$\Delta^{(n)} = \text{diag}_{p \in \mathbf{P}} W_p^{(n)} \otimes I_{\mathcal{H}_p^\circ} \quad \text{with} \quad \|\Delta^{(n)}\| \leq 1, \quad (4.16)$$

along with unit vectors $q^{(n)} \in \ell^2 \otimes \mathcal{H}_{\mathbf{S}}^\circ = \bigoplus_{p \in \mathbf{P}} \ell^2 \otimes \tilde{\mathcal{H}}_{\mathbf{S}_p} \otimes \mathcal{H}_p^\circ$ so that

$$\|(I - \Delta^{(n)} M) q^{(n)}\| < \epsilon_n. \quad (4.17)$$

Observe that then, for any $n_0 \in \mathbb{Z}_+$,

$$\begin{aligned} \epsilon_n &> \|(I - \Delta^{(n)} M) q^{(n)}\| \\ &= \|V_{\mathbf{S}}^{n_0} (I - \Delta^{(n)} M) q^{(n)}\| \\ &= \|V_{\mathbf{S}}^{n_0} q^{(n)} - V_{\mathbf{S}}^{n_0} \Delta^{(n)} V_{\mathbf{R}}^{*n_0} V_{\mathbf{R}}^{n_0} M q^{(n)}\| \\ &= \|V_{\mathbf{S}}^{n_0} q^{(n)} - \tilde{\Delta}^{(n)} M V_{\mathbf{S}}^{n_0} q^{(n)}\| \\ &= \|(I - \tilde{\Delta}^{(n)} M) \tilde{q}^{(n)}\| \end{aligned}$$

where we have set

$$\tilde{\Delta}^{(n)} = V_{\mathbf{S}}^{n_0} \Delta^{(n)} V_{\mathbf{R}}^{*n_0}, \quad \tilde{q}^{(n)} = V_{\mathbf{S}}^{n_0} q^{(n)}$$

and we used the assumed shift-invariance property $V_{\mathbf{R}} M = M V_{\mathbf{S}}$ of M . Using the representation of $\overline{\mathcal{B}}\Delta_{\mathbf{G}^\circ}$ in (3.18), it follows that $\tilde{\Delta}^{(n)}$ is in $\overline{\mathcal{B}}\Delta_{\mathbf{G}^\circ}$, since $\Delta^{(n)}$ is in $\overline{\mathcal{B}}\Delta_{\mathbf{G}^\circ}$ (see (4.16)). Moreover, $\tilde{q}^{(n)}$ is again a unit vector, but now with support in $[n_0, \infty)$. Also $\tilde{\Delta}^{(n)}$ maps $\ell^2([n_0, \infty)) \otimes \mathcal{H}_{\mathbf{R}}^\circ$ into $\ell^2([n_0, \infty)) \otimes \mathcal{H}_{\mathbf{S}}^\circ$. We conclude that without loss of generality we may assume that (4.17) holds with the additional normalization that the unit vector $q^{(n)}$ has support in $[n_0, \infty)$ and $\Delta^{(n)} \in \overline{\mathcal{B}}\Delta_{\mathbf{G}^\circ} \cap \mathcal{L}(\ell^2([n_0, \infty)) \otimes \mathcal{H}_{\mathbf{R}}^\circ, \ell^2([n_0, \infty)) \otimes \mathcal{H}_{\mathbf{S}}^\circ)$ where n_0 is any nonnegative integer of our choosing.

A familiar fact is that $P_N \rightarrow I$ strongly as $N \rightarrow \infty$. We now develop several consequences of this observation.

From the identity

$$\begin{aligned} (I - P_N \Delta^{(n)} M) P_N q^{(n)} &= \\ (I - \Delta^{(n)} M) P_N q^{(n)} &+ (I - P_N) \Delta^{(n)} M q^{(n)} - (I - P_N) \Delta^{(n)} M (I - P_N) q^{(n)} \end{aligned}$$

we get the estimate

$$\begin{aligned} \|(I - P_N \Delta^{(n)} M) P_N q^{(n)}\| &\leq \|(I - \Delta^{(n)} M) P_N q^{(n)}\| + \|(I - P_N) \Delta^{(n)} M q^{(n)}\| \\ &\quad + \|(I - P_N) \Delta^{(n)} M (I - P_N) q^{(n)}\| \\ &\leq \|(I - \Delta^{(n)} M) P_N q^{(n)}\| + \|(I - P_N) \Delta^{(n)} M q^{(n)}\| + \|M\| \|(I - P_N) q^{(n)}\|. \end{aligned}$$

By the strong convergence of $\{P_N\}$ to the identity operator, the last two terms of the final expression tend to 0 as $N \rightarrow \infty$. We arrive at the estimate

$$\|(I - P_N \Delta^{(n)} M) P_N q^{(n)}\| < \epsilon_n \text{ for } N \text{ sufficiently large.} \quad (4.18)$$

As we are assuming that $q^{(n)}$ has support in $[n_0, \infty)$, from the shift-invariance of M and the observation made above that $\Delta^{(n)}$ preserves signals with support in $[n_0, \infty)$, we see that (4.18) can be rewritten as

$$\|(I - P_{[n_0, N)} \Delta^{(n)} M) P_{[n_0, N)} q^{(n)}\| < \epsilon \quad (4.19)$$

for N sufficiently large. Note that $\text{supp } q^{(n)} \subset [n_0, \infty)$ implies that $\text{supp } Mq^{(n)} \subset [n_0, \infty)$ since M by assumption is shift invariant. We next use the identity

$$(I - P_N)MP_Nq^{(n)} = (I - P_N)Mq^{(n)} - (I - P_N)M(I - P_N)q^{(n)}$$

to get the estimate

$$\begin{aligned} \|(I - P_N)MP_Nq^{(n)}\| &\leq \|(I - P_N)Mq^{(n)}\| + \|(I - P_N)M(I - P_N)q^{(n)}\| \\ &\leq \|(I - P_N)Mq^{(n)}\| + \|M\|(I - P_N)q^{(n)}\|. \end{aligned}$$

As another consequence of the strong convergence of P_N to the identity operator, we see that, by choosing N still larger if necessary, we may arrange that in addition to (4.19) we have

$$\|(I - P_{[n_0, N]})MP_{[n_0, N]}q^{(n)}\| < \epsilon_n. \quad (4.20)$$

Moreover, if we note that

$$\begin{aligned} &\|(I - P_{[n_0, N]})\Delta^{(n)}P_{[n_0, N]}M)P_{[n_0, N]}q^{(n)}\| \\ &\leq \|(I - P_{[n_0, N]})\Delta^{(n)}M)P_{[n_0, N]}q^{(n)}\| + \|P_{[n_0, N]}\Delta^{(n)}(I - P_{[n_0, N]})MP_{[n_0, N]}q^{(n)}\| \\ &\leq \|(I - P_{[n_0, N]})\Delta^{(n)}M)P_{[n_0, N]}q^{(n)}\| + \|(I - P_{[n_0, N]})MP_{[n_0, N]}q^{(n)}\|, \end{aligned}$$

we see as a consequence of the estimates (4.19) and (4.20) that

$$\|(I - P_{[n_0, N]})\Delta^{(n)}P_{[n_0, N]}M)P_{[n_0, N]}q^{(n)}\| < 2\epsilon_n. \quad (4.21)$$

Furthermore, by rescaling and taking N still larger if necessary, we may assume in addition that $P_{[n_0, N]}q^{(n)}$ is a unit vector. By now setting $\hat{q}^{(n)} = P_{[n_0, N]}q^{(n)}$ and $\hat{\Delta}^{(n)} = P_{[n_0, N]}\Delta P_{[n_0, N]}$, and rewriting (4.21) and (4.20) in the new notation, we arrive at the following result of all this discussion: *for each $n_0 \in \mathbb{Z}_+$, there is a choice of sufficiently large $N \in \mathbb{Z}_+$ so that the following holds true: there is a unit vector $\hat{q}^{(n)}$ in $\ell^2[n_0, N) \otimes \mathcal{H}_{\mathbf{S}}^\circ$ and an operator $\hat{\Delta}^{(n)}$ in $\overline{\mathcal{B}}\Delta_{\mathbf{G}^\circ} \cap \mathcal{L}(\ell^2[n_0, N) \otimes \mathcal{H}_{\mathbf{R}}^\circ, \ell^2[n_0, N) \otimes \mathcal{H}_{\mathbf{S}}^\circ)$ such that*

$$\|(I - \hat{\Delta}^{(n)}M)\hat{q}^{(n)}\| < 2\epsilon_n \quad \text{and} \quad \|(I - P_{[n_0, N]})M\hat{q}^{(n)}\| < \epsilon_n. \quad (4.22)$$

By proceeding inductively, we may assume furthermore that the support of $\hat{q}^{(n)}$ is in an interval of the form $[t_n, t_{n+1}) \subset \mathbb{Z}_+$ with $t_0 = 0$ in such a way that these intervals form a complete partition of \mathbb{Z}_+ . In this new notation $\hat{\Delta}^{(n)}$ is in $\overline{\mathcal{B}}\Delta_{\mathbf{G}^\circ} \cap \mathcal{L}(\ell^2[t_n, t_{n+1}) \otimes \mathcal{H}_{\mathbf{R}}^\circ, \ell^2[t_n, t_{n+1}) \otimes \mathcal{H}_{\mathbf{S}}^\circ)$. If we set $\hat{\Delta} = \sum_{n=0}^{\infty} \hat{\Delta}^{(n)}P_{[t_n, t_{n+1})}$, then $\|\hat{\Delta}\| \leq 1$ since each $\hat{\Delta}^{(n)}$ is contractive and furthermore $\hat{\Delta}$ still has the block diagonal structure to qualify as an element of $\Delta_{\mathbf{G}^\circ}$, i.e., $\hat{\Delta} \in \overline{\mathcal{B}}\Delta_{\mathbf{G}^\circ}$. We now apply $(I - \hat{\Delta}M)$ to $\hat{q}^{(n)}$ and estimate the norm of the result:

$$\begin{aligned} \|(I - \hat{\Delta}M)\hat{q}^{(n)}\| &= \|(I - \hat{\Delta}\{P_{[t_n, t_{n+1})} + (I - P_{[t_n, t_{n+1})})\})M)\hat{q}^{(n)}\| \\ &= \|(I - \hat{\Delta}^{(n)}M)\hat{q}^{(n)} - \hat{\Delta}(I - P_{[t_n, t_{n+1})})M\hat{q}^{(n)}\| \\ &\leq \|(I - \hat{\Delta}^{(n)}M)\hat{q}^{(n)}\| + \|(I - P_{[t_n, t_{n+1})})M\hat{q}^{(n)}\| \\ &< 2\epsilon_n + \epsilon_n = 3\epsilon_n \end{aligned}$$

where we used (4.22) for the last inequality. As each $\hat{q}^{(n)}$ is a unit vector and $3\epsilon_n \rightarrow 0$ as $n \rightarrow \infty$, we conclude that $I - \hat{\Delta}M$ cannot be invertible, despite the fact that $\hat{\Delta} \in \overline{\mathcal{B}}\Delta_{\mathbf{G}^\circ}$. This contradiction to our underlying hypothesis completes the proof of Lemma 4.3. \square

Remark 4.4. We have seen that the strict Bounded Real Lemma (Theorem 4.2) with the help of Lemma 4.3 implies the Main Result (Theorem 3.2). It is of interest that conversely Theorem 3.2 implies Theorem 4.2 by a simple direct argument as follows. Suppose that we are given a colligation matrix $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$ as in Theorem 4.2. By hypothesis we have

$$\|I_{\ell^2} \otimes D + (I_{\ell^2} \otimes C)(I - \Delta_1(I_{\ell^2} \otimes A))^{-1}\Delta_1(I_{\ell^2} \otimes B)\| \leq \rho < 1$$

for all $\Delta_1 \in \overline{\mathbf{B}}\Delta_{\mathbf{G}^\circ}$. By a Schur-complement argument (see [45, Theorem 11.7] known as the Main Loop Theorem), this is the same as the block 2×2 matrix $\begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} - \begin{bmatrix} \Delta_1 & 0 \\ 0 & \Delta_2 \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix}$ being invertible for all $\Delta_1 \in \overline{\mathbf{B}}\Delta_{\mathbf{G}^\circ}$ and $\Delta_2 \in \overline{\mathbf{B}}\Delta_{\text{full}}$, where we set Δ_{full} equal to the set of all operators from \mathcal{U} to \mathcal{Y} . This in turn is the same as the statement

$$\mu_{\Delta_{\mathbf{G}^\circ} \oplus \Delta_{\text{full}}}(I_{\ell^2} \otimes \begin{bmatrix} A & B \\ C & D \end{bmatrix}) < 1.$$

An application of Theorem 3.2 now tells us that there exists a positive-definite matrix Γ_p° on \mathcal{H}_p for each $p \in \mathbf{P}$ and a positive real number $r > 0$ so that

$$\begin{bmatrix} A^* & C^* \\ B^* & D^* \end{bmatrix} \begin{bmatrix} \bigoplus_{r \in \mathbf{R}} \Gamma_{[r]}^\circ & 0 \\ 0 & sI_{\mathcal{Y}} \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} - \begin{bmatrix} \bigoplus_{s \in \mathbf{S}} \Gamma_{[s]}^\circ & 0 \\ 0 & sI_{\mathcal{U}} \end{bmatrix} \prec 0.$$

If we divide out by the positive numbers s and replace Γ_p° by $\frac{1}{s} \cdot \Gamma_p^\circ$ for each $p \in \mathbf{P}$, we arrive at exactly statement (3) in Theorem 4.2.

This analysis can be taken one step further to get a new proof of the strict version of the realization result Theorem 4.1 as follows. Given a rational formal power series in the strict Schur-Agler class, using results from [9] (closely related to the much earlier realization results of Fliess [23]), one can obtain a finite-dimensional colligation matrix \mathbf{U} as in (4.5) giving rise to a realization (4.6) for $S(z)$. Then use the strict Bounded Real Lemma (which as we have just seen is a direct consequence of the $\tilde{\mu} = \hat{\mu}$ result Theorem 3.2) to obtain a structured state-space similarity transforming the colligation matrix $\mathbf{U} = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$ to the strictly contractive colligation matrix $\mathbf{U}' = \begin{bmatrix} A' & B' \\ C' & D' \end{bmatrix}$. Then \mathbf{U}' is a strictly contractive colligation matrix with transfer function (4.6) (with \mathbf{U}' in place of \mathbf{U}) equal to $S(z)$, and the strict version of Theorem 4.1 follows.

Remark 4.5. It is possible to note now that Theorem 3.2 cannot be true if any of the partial state spaces \mathcal{H}_p is allowed to be infinite-dimensional and/or if the graph \mathbf{G} is allowed to be infinite. Indeed it is known (see [7]) that the Bounded Real Lemma fails if the state space is allowed to be infinite-dimensional; the proof relies on the State Space Similarity Theorem which in turn only guarantees a possibly unbounded pseudo-similarity in the infinite-dimensional setting rather than a properly bounded and boundedly invertible similarity. A simple adaptation of the example given in [7] shows that the strict Bounded Real Lemma also fails in the case of infinite-dimensional state space as well. By the preceding Remark 4.4, Theorem 3.2 is equivalent to the Bounded Real Lemma in the free noncommutative setting. We conclude that Theorem 3.2 cannot hold in general when \mathcal{H}_p is allowed to be infinite-dimensional or if the graph \mathbf{G} is allowed to have infinitely many connected components.

It is interesting to note however that Lemma 4.3 apparently does not require the finite-dimensionality of the coefficient spaces $\mathcal{H}_{\mathbf{R}}^\circ$ and $\mathcal{H}_{\mathbf{S}}^\circ$; one only requires that the operator M be shift-invariant with respect to the pair of shifts $(V_{\mathbf{S}}, V_{\mathbf{R}})$, even possibly of infinite multiplicity.

In our second proof of Theorem 3.2 in Section 5, the reader will notice several places where the finite-dimensionality of the coefficient spaces \mathcal{H}_p° and the finiteness of the graph are used—see in particular the assumed equivalence of Hilbert-Schmidt norm and operator norm in the verification of Step 1 and in the estimate (5.26).

Remark 4.6. In the *graded version* of the structured ball

$$\mathcal{B}\Delta_{\mathbf{G}^\circ} = \{L_{\mathbf{G}^\circ}(Z) = \sum_{e \in \mathbf{E}} Z_e \otimes L_{\mathbf{G}^\circ, e} : Z_e \in \mathcal{L}(\mathcal{K}), \|L_{\mathbf{G}^\circ}(Z)\| < 1\},$$

one restricts Z_e to finite square matrices $Z_e \in \mathbb{C}^{n \times n}$ for every matrix size $n = 1, 2, \dots$ rather than letting Z_e range over all bounded linear operators on a fixed infinite-dimensional separable Hilbert space \mathcal{K} . The preimage of this graded structured ball under the pencil, namely

$$\mathcal{B}\Delta_{\mathbf{G}^\circ}^{\text{pre,graded}} = \{Z = (Z_e)_{e \in \mathbf{E}} : Z_e \in \mathbb{C}^{n \times n} \text{ for } n = 1, 2, \dots, \|L_{\mathbf{G}^\circ}(Z)\| < 1\}$$

corresponds to the noncommutative pencil ball studied by Helton, Klep, McCullough and Slingend in [25, 26]. Actually these authors consider the more general setting where the formal pencil $L_{\mathbf{G}^\circ}(z)$ is replaced by a general formal pencil $L(z) = \sum_{e \in \mathbf{E}} L_e z_e$ where here \mathbf{E} is now just a convenient index set and the coefficients L_e no longer have any connection with an underlying graph. More generally, Agler and McCarthy [2] obtained a graded version of the realization result Theorem 4.1, where the structured matrix pencil $L_{\mathbf{G}^\circ}(z)$ is replaced by an arbitrary formal polynomial $\delta(z)$ with matrix coefficients, thereby obtaining a graded noncommutative analogue of the commutative result of Ball-Bolotnikov [8] and Ambrozie-Timotin [5]. This more general formalism led to new results on polynomial approximation and rigidity results for proper analytic maps between such domains, respectively for the noncommutative setting. We point out here, however, that when one replaces the structure noncommutative pencil $L_{\mathbf{G}^\circ}(z)$ by a general noncommutative pencil $L(z)$ or a general matrix noncommutative polynomial $\delta(z)$, one loses other results involving the more detailed structure of the associated noncommutative linear systems, specifically, the State Space Similarity Theorem from [9] and hence also the Bounded Real Lemma from [11] (Theorem 4.2).

5. NONCOMMUTATIVE STRUCTURED SINGULAR VALUE VERSUS DIAGONAL SCALING: A DIRECT CONVEXITY ARGUMENT FOR THE HIGHER MULTIPLICITY CASE

In this section we present our second proof of the Main Result (Theorem 3.2), this time based on the convexity-analysis approach of Dullerud and Paganini [35, 22]. In fact the approach enables one to prove the following more general formulation of Theorem 3.2. Note that Theorem 3.2 follows from the following Theorem 5.1 by setting $M = I_{\ell^2} \otimes M^\circ$.

Theorem 5.1. *Let $\overline{\mathbf{G}}$, \mathbf{G} , and \mathbf{G}° be as in Section 3.3 with $\mathcal{K} = \ell^2$, let M be a linear operator from the space*

$$\mathcal{H}_{\mathbf{S}} = \ell^2 \otimes \mathcal{H}_{\mathbf{S}}^\circ = \bigoplus_{p \in \mathbf{P}} \left(\ell^2 \otimes \tilde{\mathcal{H}}_{\mathbf{S}, p} \otimes \mathcal{H}_p^\circ \right)$$

to the space

$$\mathcal{H}_{\mathbf{R}} = \ell^2 \otimes \mathcal{H}_{\mathbf{R}}^\circ = \bigoplus_{p \in \mathbf{P}} \left(\ell^2 \otimes \tilde{\mathcal{H}}_{\mathbf{R}, p} \otimes \mathcal{H}_p^\circ \right)$$

which is shift-invariant:

$$V_{\mathbf{R}}M = MV_{\mathbf{S}} \quad \text{where} \quad V_{\mathbf{R}} = V \otimes I_{\mathcal{H}_{\mathbf{S}}^{\circ}}, \quad V_{\mathbf{S}} = V \otimes I_{\mathcal{H}_{\mathbf{R}}^{\circ}}$$

with V is the unilateral shift operator on ℓ^2 . Assume also that

- (i) the graph \mathbf{G}° has only finitely many components, and
- (ii) each coefficient space \mathcal{H}_p° is finite-dimensional.

Then

$$\mu_{\Delta_{\overline{\mathbf{G}}}}(M) = \widehat{\mu}_{\Delta_{\overline{\mathbf{G}}}}(M).$$

In particular, the following conditions are equivalent:

- (1) $\mu_{\Delta_{\overline{\mathbf{G}}}}(M) < 1$, i.e., $(I - \Delta M)^{-1}$ exists for all $\Delta = \text{diag}_{p \in \mathbf{P}} W_p \otimes I_{\mathcal{H}_p^{\circ}}$ with $W_p \in \rho \overline{\mathcal{BL}}(\ell^2 \otimes \widetilde{\mathcal{H}}_{\mathbf{R},p}, \ell^2 \otimes \widetilde{\mathcal{H}}_{\mathbf{S},p})$ for some $\rho > 1$.
- (2) $\widehat{\mu}_{\Delta_{\overline{\mathbf{G}}}}(M) < 1$, i.e., there exists operators $\Gamma_p^{\circ} \succ 0$ on \mathcal{H}_p° so that

$$M^* \left(\bigoplus_{p \in \mathbf{P}} I_{\ell^2 \otimes \widetilde{\mathcal{H}}_{\mathbf{R},p}} \otimes \Gamma_p^{\circ} \right) M - \left(\bigoplus_{p \in \mathbf{P}} I_{\ell^2 \otimes \widetilde{\mathcal{H}}_{\mathbf{S},p}} \otimes \Gamma_p^{\circ} \right) \prec 0. \quad (5.1)$$

Proof. Following the argumentation in Remark 3.3, a scaling argument gives that the equality $\mu_{\Delta_{\overline{\mathbf{G}}}}(M) = \widehat{\mu}_{\Delta_{\overline{\mathbf{G}}}}(M)$ is equivalent to: $\mu_{\Delta_{\overline{\mathbf{G}}}}(M) < 1 \Leftrightarrow \widehat{\mu}_{\Delta_{\overline{\mathbf{G}}}}(M) < 1$. This latter statement in turn is equivalent to the equivalence of the two statements (1) and (2) in the statement of the theorem. Thus it suffices to show the equivalence of (1) and (2). If (2) holds, then M is $\overline{\mathbf{G}}$ -structured-similar to a strict contraction M' from which (1) follows. We conclude that it suffices to show that (1) \Rightarrow (2).

Toward this goal, we assume that we are given M for which (1) holds. Let us use the short-hand notation

$$U_{\mathbf{R},p} = U_{\ell^2 \otimes \widetilde{\mathcal{H}}_{\mathbf{R},p}, \mathcal{H}_p^{\circ}}, \quad U_{\mathbf{S},p} = U_{\ell^2 \otimes \widetilde{\mathcal{H}}_{\mathbf{S},p}, \mathcal{H}_p^{\circ}} \quad (5.2)$$

for the identification maps between tensor product spaces and Hilbert-Schmidt operators given in Proposition 2.1. We introduce maps $\phi_p: \mathcal{H}_{\mathbf{S}} \rightarrow \mathcal{C}_1(\mathcal{H}_p^{\circ})$ by

$$\phi_p: h \mapsto (U_{\mathbf{R},p}[P_{\mathcal{H}_{\mathbf{R},p}} M h])^* U_{\mathbf{R},p}[P_{\mathcal{H}_{\mathbf{R},p}} M h] - (U_{\mathbf{S},p}[P_{\mathcal{H}_{\mathbf{S},p}} h])^* U_{\mathbf{S},p}[P_{\mathcal{H}_{\mathbf{S},p}} h] \quad (5.3)$$

In addition introduce sets of operator tuples

$$\nabla = \{(\phi_p(h))_{p \in \mathbf{P}} : h \in \mathcal{H}_{\mathbf{S}}, \|h\| = 1\}, \quad (5.4)$$

$$\Pi = \{(L_p)_{p \in \mathbf{P}} : L_p \in \mathcal{C}_1(\mathcal{H}_p^{\circ}), L_p \succeq 0, p \in \mathbf{P}\}. \quad (5.5)$$

The connection between the quadratic forms ϕ_p and the condition $\widehat{\mu}(M) < 1$ (condition (1) in Theorem 5.1) is as follows.

Lemma 5.2. *Assume that $\mu_{\Delta_{\overline{\mathbf{G}}}}(M) < 1$ (i.e., condition (1) in Theorem 5.1 is satisfied). Then*

$$\nabla \cap \Pi = \emptyset, \quad (5.6)$$

i.e., there cannot exist a nonzero $h \in \mathcal{H}_{\mathbf{S}}$ such that $\phi_p(h) \succeq 0$ for each $p \in \mathbf{P}$.

Proof. First note that each ϕ_p is homogeneous of degree 2: $\phi_p(\alpha h) = |\alpha|^2 \phi_p(h)$ for $\alpha \in \mathbb{C}$. Thus the existence of a nonzero $h \in \mathcal{H}_{\mathbf{S}}$ with $\phi_p(h) \succeq 0$ for all p implies that the normalization $\widetilde{h} = \|h\|^{-1} h$ of h is a unit vector which satisfies $\phi_p(\widetilde{h}) \succeq 0$ for all p . Thus the existence of a nonzero $h \in \mathcal{H}_{\mathbf{S}}$ with $\phi_p(h) \succeq 0$ for all p is equivalent to $\nabla \cap \Pi$ being nonempty.

To prove the lemma we proceed by contradiction. Suppose that there is a nonzero $h \in \mathcal{H}_{\mathbf{S}}$ such that $\phi_p(h) \geq 0$ for all p . By Proposition 2.3 we can find a contraction $W_p \in \mathcal{L}(\ell^2 \otimes \tilde{\mathcal{H}}_{\mathbf{R},p}, \ell^2 \otimes \tilde{\mathcal{H}}_{\mathbf{S},p})$ so that

$$W_p \otimes I_{\mathcal{H}_p^\circ} : P_{\mathcal{H}_{\mathbf{R},p}} Mh \rightarrow P_{\mathcal{H}_{\mathbf{S},p}} h.$$

Then $\Delta = \bigoplus_{p \in \mathbf{P}} (W_p \otimes I_{\mathcal{H}_p^\circ})$ is in $\overline{\mathcal{B}}\Delta_{\overline{\mathbf{G}}}$ and h is in the kernel of $(I - \Delta M)$. It follows that $I - \Delta M$ is not invertible, i.e., condition (1) in Theorem 5.1 is violated. \square

The connection of the quadratic forms ϕ_p with the condition $\hat{\mu}_{\Delta_{\mathbf{G}}}(M) < 1$ (condition (2) in Theorem 5.1) is as follows.

Lemma 5.3. *The condition $\hat{\mu}_{\Delta_{\mathbf{G}}}(M) < 1$ holds, i.e., for each $p \in \mathbf{P}$ there exists $\Gamma_p^\circ \succ 0$ on \mathcal{H}_p° so that (5.1) holds, if and only if either of the following two equivalent conditions holds:*

- (1) *There exists $\epsilon > 0$ and strictly positive definite operators $\Gamma_p^\circ \succ 0$ on \mathcal{H}_p° for each $p \in \mathbf{P}$ so that*

$$\sum_{p \in \mathbf{P}} \text{tr}(\Gamma_p^\circ \phi_p(h)) \leq -\epsilon \|h\|^2 \quad (h \in \mathcal{H}_{\mathbf{S}}). \quad (5.7)$$

- (2) *The sets ∇ and Π are strictly separated in the following sense: there exists operators Γ_p° on \mathcal{H}_p° for $p \in \mathbf{P}$ and real numbers $\alpha < \beta$ so that*

$$\sum_{p \in \mathbf{P}} \text{Re tr}(\Gamma_p^\circ K_p) \leq \alpha < \beta \leq \sum_{p \in \mathbf{P}} \text{Re tr}(\Gamma_p^\circ L_p) \quad ((K_p)_{p \in \mathbf{P}} \in \nabla, (L_p)_{p \in \mathbf{P}} \in \Pi). \quad (5.8)$$

Furthermore, whenever this is the case, it can be arranged that $\beta = 0$ and $\Gamma_p^\circ \succ 0$ and then (5.8) can be written without the real-part qualifier:

$$\sum_{p \in \mathbf{P}} \text{tr}(\Gamma_p^\circ K_p) \leq \alpha < 0 \leq \sum_{p \in \mathbf{P}} \text{tr}(\Gamma_p^\circ L_p) \quad ((K_p)_{p \in \mathbf{P}} \in \nabla, (L_p)_{p \in \mathbf{P}} \in \Pi). \quad (5.9)$$

Proof. Rewrite (5.1) as a quadratic form condition:

$$\left\langle \left[M^* \left(\bigoplus_{p \in \mathbf{P}} I_{\ell^2 \otimes \tilde{\mathcal{H}}_{\mathbf{R},p}} \otimes \Gamma_p^\circ \right) M - \left(\bigoplus_{p \in \mathbf{P}} I_{\ell^2 \otimes \tilde{\mathcal{H}}_{\mathbf{S},p}} \otimes \Gamma_p^\circ \right) \right] h, h \right\rangle \leq -\epsilon \|h\|^2. \quad (5.10)$$

The left-hand side of this inequality can be rewritten as a difference of sums:

$$\begin{aligned} & \left\langle \left[M^* \left(\bigoplus_{p \in \mathbf{P}} I_{\ell^2 \otimes \tilde{\mathcal{H}}_{\mathbf{R},p}} \otimes \Gamma_p^\circ \right) M - \left(\bigoplus_{p \in \mathbf{P}} I_{\ell^2 \otimes \tilde{\mathcal{H}}_{\mathbf{S},p}} \otimes \Gamma_p^\circ \right) \right] h, h \right\rangle \\ &= \sum_{p \in \mathbf{P}} \left\langle \left(I_{\ell^2 \otimes \tilde{\mathcal{H}}_{\mathbf{R},p}} \otimes \Gamma_p^\circ \right) P_{\mathcal{H}_{\mathbf{R},p}} Mh, P_{\mathcal{H}_{\mathbf{R},p}} Mh \right\rangle_{\mathcal{H}_{\mathbf{R},p}} \\ & \quad - \sum_{p \in \mathbf{P}} \left\langle \left(I_{\ell^2 \otimes \tilde{\mathcal{H}}_{\mathbf{S},p}} \otimes \Gamma_p^\circ \right) P_{\mathcal{H}_{\mathbf{S},p}} h, P_{\mathcal{H}_{\mathbf{S},p}} h \right\rangle_{\mathcal{H}_{\mathbf{S},p}} \end{aligned}$$

Now note that

$$\begin{aligned}
& \left\langle \left(I_{\ell^2 \otimes \tilde{\mathcal{H}}_{\mathbf{R},p}} \otimes \Gamma_p^\circ \right) P_{\mathcal{H}_{\mathbf{R},p}} Mh, P_{\mathcal{H}_{\mathbf{R},p}} Mh \right\rangle \\
&= \left\langle U_{\ell^2 \otimes \mathcal{H}_{\mathbf{R},p}, \mathcal{H}_p^\circ} \left[\left(I_{\ell^2 \otimes \tilde{\mathcal{H}}_{\mathbf{R},p}} \otimes \Gamma_p^\circ \right) P_{\mathcal{H}_{\mathbf{R},p}} Mh \right], U_{\ell^2 \otimes \tilde{\mathcal{H}}_{\mathbf{R},p}, \mathcal{H}_p^\circ} [P_{\mathcal{H}_{\mathbf{R},p}} Mh] \right\rangle \\
&= \left\langle U_{\ell^2 \otimes \mathcal{H}_{\mathbf{R},p}, \mathcal{H}_p^\circ} [P_{\mathcal{H}_{\mathbf{R},p}} Mh] (\Gamma_p^\circ)^T, U_{\ell^2 \otimes \tilde{\mathcal{H}}_{\mathbf{R},p}, \mathcal{H}_p^\circ} [P_{\mathcal{H}_{\mathbf{R},p}} Mh] \right\rangle \text{ (by property (2.3))} \\
&= \text{tr} \left((\Gamma_p^\circ)^T U_{\ell^2 \otimes \mathcal{H}_{\mathbf{R},p}, \mathcal{H}_p^\circ} [P_{\mathcal{H}_{\mathbf{R},p}} Mh]^* U_{\ell^2 \otimes \tilde{\mathcal{H}}_{\mathbf{R},p}, \mathcal{H}_p^\circ} [P_{\mathcal{H}_{\mathbf{R},p}} Mh] \right)
\end{aligned}$$

A similar calculation gives that

$$\begin{aligned}
& \left\langle \left(I_{\ell^2 \otimes \tilde{\mathcal{H}}_{\mathbf{S},p}} \otimes \Gamma_p^\circ \right) P_{\mathcal{H}_{\mathbf{S},p}} h, P_{\mathcal{H}_{\mathbf{S},p}} h \right\rangle_{\mathcal{H}_{\mathbf{S},p}} \\
&= \text{tr} \left((\Gamma_p^\circ)^T U_{\ell^2 \otimes \mathcal{H}_{\mathbf{S},p}, \mathcal{H}_p^\circ} [P_{\mathcal{H}_{\mathbf{S},p}} h]^* U_{\ell^2 \otimes \tilde{\mathcal{H}}_{\mathbf{S},p}, \mathcal{H}_p^\circ} [P_{\mathcal{H}_{\mathbf{S},p}} h] \right)
\end{aligned}$$

Putting the pieces together, we see that the condition (5.10) collapses to (5.7) (with $(\Gamma_p^\circ)^T$ in place of Γ_p°). Since the conjugation operator preserves strict positive-definiteness and is involutive, having $(\Gamma_p^\circ)^T$ in the formula rather than Γ_p° does not affect the result. Conversely, by reversing the steps in the argument, one can derive (5.1) from (5.7). This completes the proof of the equivalence of (5.1) and (5.7).

It remains to argue the equivalence of conditions (1) and (2) in Lemma 5.3. Assume that condition (1) holds, i.e., that there are positive definite operators Γ_p° on \mathcal{H}_p° for each $p \in \mathbf{P}$ so that (5.7) holds. Each $(K_p)_{p \in \mathbf{P}}$ in ∇ has the form $K_p = \phi_p(h)$ for an $h \in \mathcal{H}_{\mathbf{S}}$ with $\|h\|_{\mathcal{H}_{\mathbf{S}}} = 1$. Using this connection between $(K_p)_{p \in \mathbf{P}}$ in ∇ and h in formula (5.7) gives

$$\sum_{p \in \mathbf{P}} \text{tr}(\Gamma_p^\circ K_p) = \sum_{p \in \mathbf{P}} \text{tr}(\Gamma_p^\circ \phi_p(h)) \leq -\epsilon =: \alpha < 0.$$

Furthermore, for $\Gamma_p^\circ \succ 0$ and $L_p \succeq 0$, it is automatic that $\text{tr}(\Gamma_p^\circ L_p) \geq 0$ for each p , and hence (5.8) follows with $\beta = 0$ (and all $\Gamma_p^\circ \succ 0$).

Conversely, suppose that there are operators Γ_p° on \mathcal{H}_p° and numbers $\alpha < \beta$ so that (5.8) holds. As in general $\text{tr}(X^*) = \overline{\text{tr}(X)}$ and all components of elements of ∇ and of Π are selfadjoint, we see that $((\Gamma_p^\circ)^*)_{p \in \mathbf{P}}$ satisfies (5.8) whenever $(\Gamma_p^\circ)_{p \in \mathbf{P}}$ does. By the convexity of the conditions in (5.8), we may replace each Γ_p° by $\text{Re}\Gamma_p^\circ = \frac{1}{2}(\Gamma_p^\circ + (\Gamma_p^\circ)^*)$ and still have a solution of (5.8). Once this is done, then the presence of the real-part symbol in the formula is redundant and may be removed. At this stage we know: each Γ_p° is selfadjoint and

$$\sum_{p \in \mathbf{P}} \text{tr}(\Gamma_p^\circ K_p) \leq \alpha < \beta \leq \sum_{p \in \mathbf{P}} \text{tr}(\Gamma_p^\circ L_p) \text{ for } (K_p)_{p \in \mathbf{P}} \in \nabla \text{ and } (L_p)_{p \in \mathbf{P}} \in \Pi. \tag{5.11}$$

A particular consequence of (5.11) is that

$$\beta \leq \sum_{p \in \mathbf{P}} \text{tr}(\Gamma_p^\circ L_p) \quad \text{for} \quad L_p \succeq 0.$$

Fix a $p_o \in \mathbf{P}$ and apply this condition to the particular case where $L_p = 0$ for all $p \neq p_o$. Then we see that $\text{tr}(\Gamma_{p_o}^\circ L_{p_o}) \geq 0$ for all $L_{p_o} \succeq 0$ on $\mathcal{H}_{p_o}^\circ$. Apply this condition to the particular case where $L_{p_o} = vv^*$ for a unit vector $v \in \mathcal{H}_{p_o}^\circ$. If it were not the case that $\langle \Gamma_{p_o}^\circ v, v \rangle = \text{tr}(\Gamma_{p_o}^\circ L_{p_o}) \geq 0$, then we could rescale v to make $\text{tr}(\Gamma_{p_o}^\circ L_{p_o})$ tend as close as we like to $-\infty$, in particular, to achieve a value strictly

less than β in violation of condition (5.11). We conclude that $\Gamma_{p_0}^\circ \succeq 0$ for each $p_0 \in \mathbf{P}$ and that there is no loss of generality in taking $\beta = 0$ and then $\alpha < 0$.

It remains to see that $\Gamma_p^\circ \succ 0$ for each $p \in \mathbf{P}$. Toward this end, note that ∇ is a bounded subset of $(\mathcal{C}_1(\mathcal{H}_p^\circ))_{p \in \mathbf{P}}$ and $\alpha < 0$. Hence we may perturb each Γ_p° to $\Gamma_p^\circ + \delta I_{\mathcal{H}_p^\circ}$ for some number $\delta > 0$ sufficiently small while maintaining $\text{tr}(\Gamma_p^\circ K_p) \leq \alpha' = \alpha/2 < 0$. With these adjustments, we arrive at the existence of adjusted $\Gamma_p^\circ \succ 0$ and adjusted $\alpha < 0$ so that (5.9) holds.

Once we have the validity of (5.9), it is a simple matter to make the substitution $K_p = \phi_p(h)$ with h equal to a unit vector in $\mathcal{H}_{\mathbf{S}}$ to arrive at (5.7) for the case where h is a unit vector. As both sides of (5.7) are quadratic in rescalings of h , the general case of (5.7) now follows as well. \square

The results of Lemma 5.2 and 5.3 show the apparent gap between the conditions $\mu_{\Delta_{\overline{\nabla}}}(M) < 1$ ($\nabla \cap \Pi = \emptyset$) and $\hat{\mu}_{\Delta_{\overline{\nabla}}}(M) < 1$ (∇ and Π strictly separated). In fact, the strict separation condition (5.8) says much more, namely:

$$\overline{\text{co}} \nabla \cap \Pi = \emptyset \quad (5.12)$$

where $\overline{\text{co}} \nabla$ is the closed convex hull of the set ∇ . This suggests some elementary convexity analysis. We view $\mathcal{X} := (\mathcal{C}_1(\mathcal{H}_p^\circ))_{p \in \mathbf{P}}$ as a linear topological vector space (a bit of an overblown statement since we are assuming that it is finite dimensional) with dual space viewed as operator tuples $\mathcal{X}^* := (\mathcal{L}(\mathcal{H}_p^\circ))_{p \in \mathbf{P}}$ with duality pairing given via the trace:

$$\langle (\Gamma_p^\circ)_{p \in \mathbf{P}}, (T_p)_{p \in \mathbf{P}} \rangle = \sum_{p \in \mathbf{P}} \text{tr}(\Gamma_p^\circ T_p)$$

where

$$(\Gamma_p^\circ)_{p \in \mathbf{P}} \in (\mathcal{L}(\mathcal{H}_p^\circ))_{p \in \mathbf{P}}, \quad (T_p)_{p \in \mathbf{P}} \in (\mathcal{C}_1(\mathcal{H}_p^\circ))_{p \in \mathbf{P}}.$$

Note that Π and $\overline{\text{co}} \nabla$ are closed convex sets in \mathcal{X} and furthermore, as $\overline{\text{co}} \nabla$ is closed and bounded in the finite-dimensional Banach space \mathcal{X} , $\overline{\text{co}} \nabla$ is also compact. We may therefore apply a Hahn-Banach separation theorem (see Theorem 3.4 part (b) in [39]) to conclude: *if $\overline{\text{co}} \nabla \cap \Pi = \emptyset$, then $\overline{\text{co}} \nabla$ and Π (and hence also ∇ and Π) are strictly separated*. It then follows as a consequence of Lemma 5.3 that $\hat{\mu}_{\Delta_{\overline{\nabla}}}(M) < 1$. Hence to complete the proof of Theorem 5.1, it remains only to show:

$$\mu_{\Delta_{\overline{\nabla}}}(M) < 1 \quad \Rightarrow \quad \overline{\text{co}} \nabla \cap \Pi = \emptyset. \quad (5.13)$$

The verification of the implication (5.13) proceeds in two steps:

Step 1: If $\mu_{\Delta_{\overline{\nabla}}}(M) < 1$, then the necessary condition (5.6) holds in the stronger form

$$\overline{\nabla} \cap \Pi = \emptyset. \quad (5.14)$$

Step 2: If $\mu_{\Delta_{\overline{\nabla}}}(M) < 1$, then the closure $\overline{\nabla}$ of ∇ is convex, i.e.,

$$\overline{\text{co}} \nabla = \overline{\nabla}. \quad (5.15)$$

Verification of Step 1: The argument is modeled on the proof of Lemma B.1 in [22] inspired in turn by the earlier work of Shamma [41]; the reader will see that the argument also has some elements in common with the proof of Lemma 4.3 above.

To streamline the proof, let us use the short-hand notation (5.2). Recall that V denotes the shift operator on ℓ^2 ; let us introduce additional short-hand notation

for the higher-multiplicity shift operators

$$\begin{aligned}\tilde{V}_{\mathbf{R}} &:= V \otimes I_{\tilde{\mathcal{H}}_{\mathbf{R},p}} \text{ on } \ell^2 \otimes \tilde{\mathcal{H}}_{\mathbf{R},p}, & \tilde{V}_{\mathbf{S}} &:= V \otimes I_{\tilde{\mathcal{H}}_{\mathbf{S},p}} \text{ on } \ell^2 \otimes \tilde{\mathcal{H}}_{\mathbf{S},p}, \\ V_{\mathbf{R}} &:= V \otimes I_{\mathcal{H}_{\mathbf{R},p}^\circ} \text{ on } \ell^2 \otimes \mathcal{H}_{\mathbf{R},p}^\circ, & V_{\mathbf{S}} &:= V \otimes I_{\mathcal{H}_{\mathbf{S},p}^\circ} \text{ on } \ell^2 \otimes \mathcal{H}_{\mathbf{S},p}^\circ\end{aligned}$$

where we recall that $\mathcal{H}_{\mathbf{R},p}^\circ = \tilde{\mathcal{H}}_{\mathbf{R},p} \otimes \mathcal{H}_p^\circ$ and $\mathcal{H}_{\mathbf{S},p}^\circ = \tilde{\mathcal{H}}_{\mathbf{S},p} \otimes \mathcal{H}_p^\circ$. For brevity we use the same notation for the case where $\mathcal{H}_{\mathbf{S},p}$ is replaced by $\mathcal{H}_{\mathbf{S}}$ or $\mathcal{H}_{\mathbf{R},p}$ is replaced by $\mathcal{H}_{\mathbf{R}}$; the meaning will be clear from the context. Then we have the identities

$$\tilde{V}_{\mathbf{R}} U_{\mathbf{R},p}[h_{\mathbf{R}}] = U_{\mathbf{R},p}[V_{\mathbf{R},p} h_{\mathbf{R}}], \quad \tilde{V}_{\mathbf{S}} U_{\mathbf{S},p}[h_{\mathbf{S}}] = U_{\mathbf{S},p}[V_{\mathbf{S}} h_{\mathbf{S}}] \quad (5.16)$$

for $h_{\mathbf{R}} \in \mathcal{H}_{\mathbf{R},p}$ and $h_{\mathbf{S}} \in \mathcal{H}_{\mathbf{S},p}$ as a consequence of property (2.3) in Proposition 2.2.

Note that the condition (5.14) can otherwise be formulated as $\text{dist}(\nabla \cap \Pi) > 0$ where the distance can be measured via any convenient norm on the trace-class operator-tuples $(\mathcal{C}_1(\mathcal{H}_p^\circ))_{p \in \mathbf{P}}$; note that as part of our assumptions is that $\dim \mathcal{H}_p^\circ < \infty$ for each p , all the norms on \mathcal{H}_p° are equivalent. For convenience we work with the operator norm.

We proceed by contradiction. Suppose that $\text{dist}(\nabla, \Pi) = 0$. The idea is to construct a $\Delta \in \overline{\mathcal{B}\Delta\overline{\mathcal{G}}}$ so that $I - \Delta M$ is not (boundedly) invertible. Let $\{\epsilon_n\}_{n \in \mathbb{Z}_+}$ be a sequence of positive real numbers with $\epsilon_n \rightarrow 0$ as $n \rightarrow \infty$. Since $\text{dist}(\nabla, \Pi) = 0$, we can find a unit vector $q^{(n)} \in \mathcal{H}_{\mathbf{S}}$ and an operator-tuple $(L_p^{(n)})_{p \in \mathbf{P}} \in \Pi$ so that

$$\|\phi_p(q^{(n)}) - L_p^{(n)}\|_{\mathcal{L}(\mathcal{H}_p^\circ)} < \epsilon_n^2 \text{ for each } p \in \mathbf{P}.$$

Since $\phi_p(q^{(n)})$ and $L_p^{(n)}$ are all selfadjoint, this norm inequality implies the quadratic form inequality

$$-\epsilon_n^2 I_{\mathcal{H}_p^\circ} \prec \phi_p(q^{(n)}) - L_p^{(n)} \preceq \epsilon_n^2 I_{\mathcal{H}_p^\circ}.$$

In particular we have

$$\phi_p(q^{(n)}) \succ L_p^{(n)} - \epsilon_n^2 I_{\mathcal{H}_p^\circ} \succeq -\epsilon_n^2 I_{\mathcal{H}_p^\circ}.$$

Spelling this condition out gives

$$\begin{aligned}& \left(U_{\mathbf{R},p}[P_{\mathcal{H}_{\mathbf{R},p}} M q^{(n)}] \right)^* U_{\mathbf{R},p}[P_{\mathcal{H}_{\mathbf{R},p}} M q^{(n)}] - \left(U_{\mathbf{S},p}[P_{\mathcal{H}_{\mathbf{S},p}} q^{(n)}] \right)^* U_{\mathbf{S},p}[P_{\mathcal{H}_{\mathbf{S},p}} q^{(n)}] \\ & \succ -\epsilon_n^2 I_{\mathcal{H}_p^\circ} \text{ for all } p \in \mathbf{P}.\end{aligned} \quad (5.17)$$

Let now n_0 be an arbitrary nonnegative integer. Note that $P_{\mathcal{H}_{\mathbf{R},p}}$ commutes with $V_{\mathbf{R},p}$. Hence, using the property (5.16) and the assumed shift-invariance of M , we see that

$$(\tilde{V}_{\mathbf{R},p})^{n_0} U_{\mathbf{R},p}[P_{\mathcal{H}_{\mathbf{R},p}} M q^{(n)}] = U_{\mathbf{R},p}[P_{\mathcal{H}_{\mathbf{R},p}} M (V_{\mathbf{S},p})^{n_0} q^{(n)}]$$

and similarly

$$(\tilde{V}_{\mathbf{S},p})^{n_0} U_{\mathbf{S},p}[P_{\mathcal{H}_{\mathbf{S},p}} q^{(n)}] = U_{\mathbf{R},p}[P_{\mathcal{H}_{\mathbf{R},p}} (V_{\mathbf{S},p})^{n_0} q^{(n)}].$$

It follows that (5.17) continues to hold with $(V_{\mathbf{S},p})^{n_0} q^{(n)}$ in place of $q^{(n)}$. Hence we may assume without loss of generality that (5.17) holds with $q^{(n)}$ a unit vector having support in $[n_0, \infty)$.

As was done in the proof of Lemma 4.3, let us write $P_{[n_0, N]}$ for the projection of $\ell^2 \otimes \mathcal{X}$ onto $\ell^2[n_0, N] \otimes \mathcal{X}$, where the coefficient space \mathcal{X} is either $\mathcal{H}_{\mathbf{S},p}^\circ = \tilde{\mathcal{H}}_{\mathbf{S},p} \otimes \mathcal{H}_p^\circ$ or $\mathcal{H}_{\mathbf{R},p}^\circ = \tilde{\mathcal{H}}_{\mathbf{R},p} \otimes \mathcal{H}_p^\circ$; in case $n_0 = 0$, we write simply P_N rather than $P_{[0, N]}$. In case the coefficient space is either $\tilde{\mathcal{H}}_{\mathbf{S},p}$ or $\tilde{\mathcal{H}}_{\mathbf{R},p}$, we write $\tilde{P}_{[n_0, N]}$ and \tilde{P}_N respectively.

The key property of the sequence $\{P_N\}_{N \in \mathbb{Z}_+}$ is its strong convergence to the identity operator as $N \rightarrow \infty$. In particular $P_N q^{(n)} \rightarrow q^{(n)}$ in \mathcal{H}_S -norm as $N \rightarrow \infty$. It follows that $U_{S,p}[P_{\mathcal{H}_{S,p}} P_N q^{(n)}] \rightarrow U_{S,p}[P_{\mathcal{H}_{S,p}} q^{(n)}]$ in Hilbert-Schmidt norm as $N \rightarrow \infty$. But since $\dim \mathcal{H}_p^\circ < \infty$, this is enough to conclude that actually $U_{S,p}[P_{\mathcal{H}_{S,p}} P_N q^{(n)}]$ converges to $U_{S,p}[P_{\mathcal{H}_{S,p}} q^{(n)}]$ in operator norm as $N \rightarrow \infty$. Similarly $U_{R,p}[P_{\mathcal{H}_{R,p}} M P_N q^{(n)}]$ converges to $U_{R,p}[P_{\mathcal{H}_{R,p}} M q^{(n)}]$ in operator norm as $N \rightarrow \infty$. We may thus arrange that (5.17) holds with $P_N q^{(n)}$ in place of $q^{(n)}$, in other words, for a fixed n_0 we may assume without loss of generality that $q^{(n)}$ has support in $[n_0, N]$ for some sufficiently large integer $N > n_0$.

From the estimate

$$\begin{aligned} & \|P_N P_{\mathcal{H}_{R,p}} M P_N q^{(n)} - P_{\mathcal{H}_{R,p}} M q^{(n)}\| \\ & \leq \|P_N P_{\mathcal{H}_{R,p}} M (P_N - I) q^{(n)}\| + \|(P_N - I) P_{\mathcal{H}_{R,p}} M q^{(n)}\| \end{aligned}$$

coupled with the strong convergence of P_N to the identity operator as $N \rightarrow \infty$, we see that $P_N P_{\mathcal{H}_{R,p}} M P_N q^{(n)}$ converges to $P_{\mathcal{H}_{R,p}} M q^{(n)}$ in $\mathcal{H}_{R,p}$ -norm. It follows that $U_{R,p}[P_N P_{\mathcal{H}_{R,p}} M P_N q^{(n)}]$ converges to $U_{R,p}[P_{\mathcal{H}_{R,p}} M q^{(n)}]$ in operator norm. Hence by taking N sufficiently large, (5.17) can be adjusted to have the form

$$\begin{aligned} & \left(U_{R,p}[P_{[n_0, N]} P_{\mathcal{H}_{R,p}} M P_{[n_0, N]} q^{(n)}] \right)^* U_{R,p}[P_{[n_0, N]} P_{\mathcal{H}_{R,p}} M P_{[n_0, N]} q^{(n)}] \\ & - \left(U_{S,p}[P_{\mathcal{H}_{S,p}} P_{[n_0, N]} q^{(n)}] \right)^* U_{S,p}[P_{\mathcal{H}_{S,p}} P_{[n_0, N]} q^{(n)}] \succ -\epsilon_n^2 I_{\mathcal{H}_p^\circ}. \end{aligned} \quad (5.18)$$

Furthermore, since $P_{[n_0, N]} M P_{[n_0, N]} = P_N M P_{[n_0, N]}$, by the shift invariance of M , we have

$$\|(I - P_{[n_0, N]}) M P_{[n_0, N]} q^{(n)}\| = \|(I - P_N) M P_{[n_0, N]} q^{(n)}\|$$

combined with the strong convergence of P_N to the identity operator as $N \rightarrow \infty$ tells us that we may also arrange that

$$\|(I - P_{[n_0, N]}) M P_{[n_0, N]} q^{(n)}\| < \epsilon_n \quad (5.19)$$

by taking N sufficiently large. By rescaling and taking N still larger if necessary, we can assume without loss of generality that $P_{[n_0, N]} q^{(n)}$ is a unit vector. By redefining $q^{(n)}$ to be $P_{[n_0, N]} q^{(n)}$, we arrive at the following normalization: *for any $n_0 \in \mathbb{Z}_+$, by taking $N \in \mathbb{Z}_+$ sufficiently large, we can find a unit vector $q^{(n)} \in \ell^2 \otimes \mathcal{H}_S^\circ$ with support in $[n_0, N]$ so that*

$$\begin{aligned} & \left(U_{R,p}[P_{[n_0, N]} P_{\mathcal{H}_{R,p}} M q^{(n)}] \right)^* U_{R,p}[P_{[n_0, N]} P_{\mathcal{H}_{R,p}} M q^{(n)}] \\ & - \left(U_{S,p}[P_{\mathcal{H}_{S,p}} q^{(n)}] \right)^* U_{S,p}[P_{\mathcal{H}_{S,p}} q^{(n)}] \succ -\epsilon_n^2 I_{\mathcal{H}_p^\circ}. \end{aligned} \quad (5.20)$$

as well as

$$\|(I - P_{[n_0, N]}) M q^{(n)}\| < \epsilon_n. \quad (5.21)$$

Let us apply the Douglas lemma to the inequality (5.20) (see the discussion immediately preceding Proposition 2.3); the result is the existence of a contraction operator

$$\begin{bmatrix} X_p^{(n)} & Y_p^{(n)} \end{bmatrix} : \begin{bmatrix} \ell^2([n_0, N]) \otimes \tilde{\mathcal{H}}_{R,p} \\ \mathcal{H}_p^\circ \end{bmatrix} \rightarrow \ell^2([n_0, N]) \otimes \tilde{\mathcal{H}}_{S,p}$$

so that

$$X_p^{(n)} U_{\mathbf{R},p} [P_{[n_0,N]} P_{\mathcal{H}_{\mathbf{R},p}} M P_{[n_0,N]} q^{(n)}] + \epsilon_n Y_p^{(n)} = U_{\mathbf{S},p} [P_{\mathcal{H}_{\mathbf{S},p}} P_{[n_0,N]} q^{(n)}].$$

Hence, as a consequence of property (2.3), we get

$$U_{\mathbf{S},p} \left[\left(X_p^{(n)} \otimes I_{\mathcal{H}_p^\circ} \right) P_{[n_0,N]} P_{\mathcal{H}_{\mathbf{R},p}} M P_{[n_0,N]} q^{(n)} - P_{\mathcal{H}_{\mathbf{S},p}} P_{[n,N]} q^{(n)} \right] = \epsilon_n Y_p^{(n)}. \quad (5.22)$$

As $\|Y_p^{(n)}\| \leq 1$, it follows that

$$\text{tr} (Y_p^{(n)*} Y_p^{(n)}) = \sum_{j \in J} \|Y_p e_j^{(p)}\|^2 \leq \dim \mathcal{H}_p^\circ$$

where we let $\{e_j^{(p)} : j \in J\}$ be any orthonormal basis for \mathcal{H}_p° . Hence we see that $Y_p^{(n)}$ has Hilbert-Schmidt norm $\|Y_p^{(n)}\|_{\mathcal{C}_2(\mathcal{H}_p^\circ, \ell^2 \otimes \tilde{\mathcal{H}}_{\mathbf{S},p})}$ at most $(\dim \mathcal{H}_p^\circ)^{1/2}$. As $U_{\mathbf{S},p}$ is unitary from $\mathcal{H}_{\mathbf{S},p}$ to $\mathcal{C}_2(\mathcal{H}_p^\circ, \ell^2 \otimes \tilde{\mathcal{H}}_{\mathbf{S},p})$, we see from equality (5.22) that

$$\|(X_p^{(n)} \otimes I_{\mathcal{H}_p^\circ}) P_{[n_0,N]} P_{\mathcal{H}_{\mathbf{R},p}} M q^{(n)} - P_{\mathcal{H}_{\mathbf{S},p}} q^{(n)}\| < \epsilon_n \cdot (\dim \mathcal{H}_p^\circ)^{1/2}. \quad (5.23)$$

If we set $\Delta^{(n)} = \text{diag}_{p \in \mathbf{P}} P_{[n_0,N]} \left(X_p^{(n)} \otimes I_{\mathcal{H}_p^\circ} \right) P_{[n_0,N]}$, then we see that $\Delta^{(n)}$ has the correct block-diagonal structure to be an element of the structure $\mathbf{\Delta}_{\overline{\mathbf{G}}}$ and furthermore $\|\Delta^{(n)}\| \leq 1$ since $\|X_p^{(n)} \otimes I_{\mathcal{H}_p^\circ}\| \leq 1$ for each p , i.e., $\Delta^{(n)} \in \overline{\mathbf{B}} \mathbf{\Delta}_{\overline{\mathbf{G}}}$. Furthermore, from (5.23) we see that

$$\|(I - \Delta^{(n)}) M q^{(n)}\| < \epsilon_n \left(\sum_{p \in \mathbf{P}} \dim \mathcal{H}_p^\circ \right)^{1/2}. \quad (5.24)$$

We now exploit the arbitrariness of n_0 in the preceding analysis. Proceeding inductively, we may take the support of $q^{(n)}$ to be contained in an interval of the form $[t_n, t_{n+1}) \subset \mathbb{Z}_+$ with $t_0 = 0$ such that these intervals form a complete partition of \mathbb{Z}_+ . Now set

$$\Delta = \sum_{n=0}^{\infty} \Delta^{(n)} P_{[t_n, t_{n+1})}.$$

Then it is easily seen that $\Delta \in \overline{\mathbf{B}} \mathbf{\Delta}_{\overline{\mathbf{G}}}$. When we apply $I - \Delta M$ to $q^{(n)}$ and estimate the norm, we get

$$\begin{aligned} \|(I - \Delta M) q^{(n)}\| &= \|(I - \Delta \{P_{[t_n, t_{n+1})} + (I - P_{[t_n, t_{n+1})})\}) M q^{(n)}\| \\ &= \|(I - \Delta^{(n)} M) q^{(n)} - \Delta (I - P_{[t_n, t_{n+1})}) M q^{(n)}\| \end{aligned} \quad (5.25)$$

$$\begin{aligned} &\leq \|(I - \Delta^{(n)} M) q^{(n)}\| + \|(I - P_{[t_n, t_{n+1})}) M q^{(n)}\| \\ &< \epsilon_n \cdot \left(\sum_{p \in \mathbf{P}} \dim \mathcal{H}_p^\circ \right)^{1/2} + \epsilon_n \end{aligned} \quad (5.26)$$

by (5.24) and (5.21). As each $q^{(n)}$ is a unit vector and $\epsilon_n \cdot (\sum_{p \in \mathbf{P}} \dim \mathcal{H}_p^\circ)^{1/2} + \epsilon_n \rightarrow 0$ as $n \rightarrow \infty$, we conclude that $I - \Delta M$ is not bounded below and hence cannot be boundedly invertible, in contradiction to the assumption that $\hat{\mu}_{\mathbf{\Delta}_{\overline{\mathbf{G}}}}(M) < 1$. This completes the verification of Step 1.

Verification of Step 2: The proof is modeled on Lemma 8.11 in [22] and follows the original idea of Megretski and Treil [33].

Let $h, \tilde{h} \in \mathcal{H}_{\mathbf{S}}$ with $\|h\| = \|\tilde{h}\| = 1$ and let $\alpha \in (0, 1)$. We shall prove that $\alpha\phi_p(h) + (1-\alpha)\phi_p(\tilde{h}) \in \overline{\nabla}$ for each $p \in \mathbf{P}$. The convexity of the set $\overline{\nabla}$ then follows via a straightforward continuity argument.

For each $n \in \mathbb{Z}_+$, set $h_n = \sqrt{\alpha}h + \sqrt{1-\alpha}V_{\mathbf{S}}^n\tilde{h} \in \mathcal{H}_{\mathbf{S}}$. Then

$$\begin{aligned}\|h_n\|^2 &= \alpha\|h\|^2 + (1-\alpha)\|V_{\mathbf{S}}^n\tilde{h}\|^2 + 2\sqrt{\alpha(1-\alpha)}\operatorname{Re}\langle h, V_{\mathbf{S}}^n\tilde{h} \rangle \\ &= \alpha + (1-\alpha) + 2\sqrt{\alpha(1-\alpha)}\operatorname{Re}\langle V_{\mathbf{S}}^{*n}h, \tilde{h} \rangle \\ &= 1 + 2\sqrt{\alpha(1-\alpha)}\operatorname{Re}\langle V_{\mathbf{S}}^{*n}h, \tilde{h} \rangle.\end{aligned}$$

Since $V_{\mathbf{S}}^{*n}$ converges strongly (hence also weakly) to 0, we conclude that

$$\|h_n\|^2 \rightarrow 1 \quad \text{as } n \rightarrow \infty. \quad (5.27)$$

Next, writing out $h_n = \sqrt{\alpha}h + \sqrt{1-\alpha}V_{\mathbf{S}}^n\tilde{h}$ and using the linearity of $U_{\mathbf{R},p}$ and $U_{\mathbf{S},p}$ we observe that

$$\begin{aligned}\phi_p(h_n) &= (U_{\mathbf{R},p}[P_{\mathcal{H}_{\mathbf{R},p}}Mh_n])^* U_{\mathbf{R},p}[P_{\mathcal{H}_{\mathbf{R},p}}Mh_n] - (U_{\mathbf{S},p}[P_{\mathcal{H}_{\mathbf{S},p}}h_n])^* U_{\mathbf{S},p}[P_{\mathcal{H}_{\mathbf{S},p}}h_n] \\ &= \alpha\phi_p(h) + (1-\alpha)\phi_p(V_{\mathbf{S}}^n\tilde{h}) + 2\sqrt{\alpha(1-\alpha)} \\ &\quad \cdot \operatorname{Re} \left((U_{\mathbf{R},p}[P_{\mathcal{H}_{\mathbf{R},p}}Mh])^* U_{\mathbf{R},p}[P_{\mathcal{H}_{\mathbf{R},p}}MV_{\mathbf{S}}^n\tilde{h}] - (U_{\mathbf{S},p}[P_{\mathcal{H}_{\mathbf{S},p}}h])^* U_{\mathbf{S},p}[P_{\mathcal{H}_{\mathbf{S},p}}V_{\mathbf{S}}^n\tilde{h}] \right).\end{aligned}$$

A consequence of the intertwining property (5.16) and the shift-invariance of M is that in fact

$$\phi_p(V_{\mathbf{S}}^n\tilde{h}) = \phi_p(\tilde{h}).$$

Thus in fact we have

$$\begin{aligned}\phi_p(h_n) &= \alpha\phi_p(h) + (1-\alpha)\phi_p(\tilde{h}) + 2\sqrt{\alpha(1-\alpha)} \\ &\quad \cdot \operatorname{Re} \left((U_{\mathbf{R},p}[P_{\mathcal{H}_{\mathbf{R},p}}Mh])^* U_{\mathbf{R},p}[P_{\mathcal{H}_{\mathbf{R},p}}MV_{\mathbf{S}}^n\tilde{h}] - (U_{\mathbf{S},p}[P_{\mathcal{H}_{\mathbf{S},p}}h])^* U_{\mathbf{S},p}[P_{\mathcal{H}_{\mathbf{S},p}}V_{\mathbf{S}}^n\tilde{h}] \right).\end{aligned} \quad (5.28)$$

We claim that the cross terms tend to zero (in trace-class norm) as $n \rightarrow \infty$. A sample term to check is

$$(U_{\mathbf{R},p}[P_{\mathcal{H}_{\mathbf{R},p}}Mh])^* U_{\mathbf{R},p}[P_{\mathcal{H}_{\mathbf{R},p}}MV_{\mathbf{S}}^n\tilde{h}] \rightarrow 0 \text{ as } n \rightarrow \infty. \quad (5.29)$$

We again use the shift invariance of M and the intertwining property (5.16) to see that

$$\begin{aligned}(U_{\mathbf{R},p}[P_{\mathcal{H}_{\mathbf{R},p}}Mh])^* U_{\mathbf{R},p}[P_{\mathcal{H}_{\mathbf{R},p}}MV_{\mathbf{S}}^n\tilde{h}] &= (U_{\mathbf{R},p}[P_{\mathcal{H}_{\mathbf{R},p}}Mh])^* \tilde{V}_{\mathbf{R}}^n U_{\mathbf{R},p}[P_{\mathcal{H}_{\mathbf{R},p}}M\tilde{h}] \\ &= \left(\tilde{V}_{\mathbf{R}}^{*n} U_{\mathbf{R},p}[P_{\mathcal{H}_{\mathbf{R},p}}Mh] \right)^* U_{\mathbf{R},p}[P_{\mathcal{H}_{\mathbf{R},p}}M\tilde{h}] \\ &= (U_{\mathbf{R},p}[V_{\mathbf{R}}^{*n} P_{\mathcal{H}_{\mathbf{R},p}}Mh])^* U_{\mathbf{R},p}[P_{\mathcal{H}_{\mathbf{R},p}}M\tilde{h}]\end{aligned}$$

The fact that $V_{\mathbf{R}}^{*n} P_{\mathcal{H}_{\mathbf{R},p}}Mh \rightarrow 0$ in $\mathcal{H}_{\mathbf{R},p}$ implies that $U_{\mathbf{R},p}[V_{\mathbf{R}}^{*n} P_{\mathcal{H}_{\mathbf{R},p}}Mh] \rightarrow 0$ in Hilbert-Schmidt norm, and hence (5.29) now follows. A similar calculation shows that

$$(U_{\mathbf{S},p}[P_{\mathcal{H}_{\mathbf{S},p}}h])^* U_{\mathbf{S},p}[P_{\mathcal{H}_{\mathbf{S},p}}V_{\mathbf{S}}^n\tilde{h}] \rightarrow 0 \text{ as } n \rightarrow \infty.$$

From (5.28) we now read off

$$\phi_p(h_n) \rightarrow \alpha\phi_p(h) + (1-\alpha)\phi_p(\tilde{h}).$$

As we observed already in (5.27) that $\|h_n\| \rightarrow 1$ as $n \rightarrow \infty$, we see that we also have

$$\phi_p \left(\frac{h_n}{\|h_n\|} \right) = \frac{1}{\|h_n\|^2} \phi(h_n) \rightarrow \alpha \phi_p(h) + (1 - \alpha) \phi_p(\tilde{h}) \text{ as } n \rightarrow \infty.$$

This exhibits the convex combination $\alpha \phi_p(h) + (1 - \alpha) \phi_p(\tilde{h})$ of two elements of ∇ as an element of $\overline{\nabla}$ and completes the verification of Step 2.

The proof of Theorem 5.1 is now complete once one observes that the combined results of Steps 1 and 2 lead immediately to the validity of the implication (5.13). \square

Remark 5.4. We note the following more general version of Theorem 3.2.

Theorem 5.5. *Let $\overline{\mathbf{G}}$ and \mathbf{G}° be as in Theorem 3.2 with the number of components of \mathbf{G}° finite and with all coefficient Hilbert spaces \mathcal{H}_p° finite-dimensional. Let \mathcal{K} be a fixed separable infinite-dimensional Hilbert space (e.g., $\mathcal{K} = \ell^2$). Then the following stabilizability and detectability results hold.*

- (1) Suppose that (A, B) is an input-pair of the form

$$\begin{bmatrix} A & B \end{bmatrix} : \begin{bmatrix} \mathcal{H}_S^\circ \\ \mathcal{U} \end{bmatrix} \rightarrow \mathcal{H}_R^\circ$$

Then the following conditions are equivalent:

- (a) The operator

$$\begin{bmatrix} I - \Delta(I \otimes A) & I \otimes B \end{bmatrix} : \begin{bmatrix} \mathcal{K} \otimes \mathcal{H}_S^\circ \\ \mathcal{K} \otimes \mathcal{U} \end{bmatrix} \rightarrow \mathcal{K} \otimes \mathcal{H}_R^\circ$$

is boundedly left invertible for all $\Delta \in \rho \overline{\mathbf{B}} \Delta_{\overline{\mathbf{G}}}$ for some $\rho > 1$.

- (b) There exist positive definite operators $\Gamma_p^\circ \succ 0$ on \mathcal{H}_p° so that

$$A \left(\bigoplus_{p \in \mathbf{P}} (I_{\mathcal{H}_{S,p}} \otimes \Gamma_p^\circ) \right) A^* - \left(\bigoplus_{p \in \mathbf{P}} I_{\mathcal{H}_{R,p}} \otimes \Gamma_p^\circ \right) - BB^* \preceq 0.$$

- (c) There exist a feedback operator $F: \mathcal{H}_S \rightarrow \mathcal{U}$ so that $\mu_{\overline{\mathbf{G}}}(I_{\mathcal{K}} \otimes (A + BF)) < 1$, i.e., for some $\rho > 1$ the operator $I - \Delta(I \otimes (A + BF))$ is boundedly invertible for each $\Delta \in \overline{\mathbf{B}} \Delta_{\overline{\mathbf{G}}}$.

- (2) Suppose that (C, A) is an output-pair of the form

$$\begin{bmatrix} A \\ C \end{bmatrix} : \mathcal{H}_S^\circ \rightarrow \begin{bmatrix} \mathcal{H}_R^\circ \\ \mathcal{Y} \end{bmatrix}.$$

Then the following conditions are equivalent:

- (a) The operator

$$\begin{bmatrix} I - \Delta(I \otimes A) \\ I \otimes C \end{bmatrix} : \mathcal{K} \otimes \mathcal{H}_S^\circ \rightarrow \begin{bmatrix} \mathcal{K} \otimes \mathcal{H}_R^\circ \\ \mathcal{K} \otimes \mathcal{Y} \end{bmatrix}$$

is boundedly left invertible for all $\Delta \in \rho \overline{\mathbf{B}} \Delta_{\overline{\mathbf{G}}}$ for some $\rho > 1$.

- (b) There exist positive definite operators $\Gamma_p^\circ \succ 0$ on \mathcal{H}_p° so that

$$A^* \left(\bigoplus_{p \in \mathbf{P}} (I_{\mathcal{H}_{R,p}} \otimes \Gamma_p^\circ) \right) A - \left(\bigoplus_{p \in \mathbf{P}} I_{\mathcal{H}_{S,p}} \otimes \Gamma_p^\circ \right) - C^* C \prec 0.$$

- (c) *There exists an output injection $L: \mathcal{Y} \rightarrow \mathcal{H}_{\mathbf{R}}$ so that $\mu_{\overline{\mathbf{G}}}(I_{\mathcal{K}} \otimes (A + LC)) < 1$, i.e., for some $\rho > 1$ the operator $I - \Delta(I \otimes (A + LC))$ is boundedly invertible for each $\Delta \in \rho \overline{\mathbf{B}} \Delta_{\overline{\mathbf{G}}}$.*

For the simple multiplicity case ($\mathcal{H}_p^\circ = \mathbb{C}$ for all p), details of this result can be found in [35]; a nice summary (with no proofs) is in [45]. We expect that either of the proofs of Theorem 3.2 presented here can be adapted to arrive at the more general formulation in Theorem 5.5. We refer also to [14] for additional information and perspective.

Remark 5.6. For the case where $M = I_{\ell^2} \otimes M^\circ$ where M° is an operator between the finite-dimensional spaces $\mathcal{H}_{\mathbf{S}}^\circ$ to $\mathcal{H}_{\mathbf{R}}^\circ$, the Linear Operator Inequality (5.1) reduces to the finite-dimensional Linear Matrix Inequality

$$M^{\circ*} \left(\bigoplus_{p \in \mathbf{P}} I_{\tilde{\mathcal{H}}_{\mathbf{R},p}} \otimes \Gamma_p^\circ \right) M^\circ - \left(\bigoplus_{p \in \mathbf{P}} I_{\tilde{\mathcal{H}}_{\mathbf{S},p}} \otimes \Gamma_p^\circ \right) \prec 0. \quad (5.30)$$

In case M is a shift-invariant operator from $\ell^2 \otimes \mathcal{H}_{\mathbf{S}}^\circ$ to $\ell^2 \otimes \mathcal{H}_{\mathbf{R}}^\circ$, there appears no reason for the LOI (5.1) to collapse to an LMI like (5.30) in general. However, if we assume that M is given via convolution with a distribution having Z -transform equal to a rational matrix function $\widehat{M}(\lambda)$ having state-space realization

$$\widehat{M}(\lambda) = D + \lambda C(I - \lambda A)^{-1} B$$

with A stable (the spectrum $\sigma(A)$ of A is inside the unit disk \mathbb{D}) where the system matrix

$$\mathbf{U} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} : \begin{bmatrix} \mathcal{X} \\ \tilde{\mathcal{H}}_{\mathbf{S}} \end{bmatrix} \rightarrow \begin{bmatrix} \mathcal{X} \\ \tilde{\mathcal{H}}_{\mathbf{R}} \end{bmatrix}$$

is finite-dimensional, then it is possible to convert the LOI (5.1) to an LMI condition as follows. Rewrite the LOI (5.1) as

$$\left\| \left(\bigoplus_{p \in \mathbf{P}} I_{\ell^2} \otimes (I_{\tilde{\mathcal{H}}_{\mathbf{R},p}} \otimes \Gamma_p^\circ)^{1/2} \right) M \left(\bigoplus_{p \in \mathbf{P}} I_{\ell^2} \otimes (I_{\tilde{\mathcal{H}}_{\mathbf{S},p}} \otimes \Gamma_p^\circ)^{1/2} \right)^{-1/2} \right\| < 1. \quad (5.31)$$

After applying the Z -transform to move to the frequency domain, we see from (5.31) that

$$\sup_{\lambda \in \mathbb{D}} \left\| \left(\bigoplus_{p \in \mathbf{P}} (I_{\tilde{\mathcal{H}}_{\mathbf{R},p}} \otimes \Gamma_p^\circ)^{1/2} \right) \widehat{M}(\lambda) \left(\bigoplus_{p \in \mathbf{P}} (I_{\tilde{\mathcal{H}}_{\mathbf{S},p}} \otimes \Gamma_p^\circ)^{-1/2} \right) \right\| < 1. \quad (5.32)$$

As we are assuming that A is stable, the standard strict Bounded Real Lemma implies that there is a positive-definite $X \succ 0$ on \mathcal{X} so such that

$$\begin{bmatrix} \tilde{A}^* & \tilde{C}^* \\ \tilde{B}^* & \tilde{D}^* \end{bmatrix} \begin{bmatrix} X & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} \tilde{A} & \tilde{B} \\ \tilde{C} & \tilde{D} \end{bmatrix} - \begin{bmatrix} X & 0 \\ 0 & I \end{bmatrix} \prec 0 \quad (5.33)$$

where we have set

$$\begin{bmatrix} \tilde{A} & \tilde{B} \\ \tilde{C} & \tilde{D} \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & (I_{\tilde{\mathcal{H}}_{\mathbf{R},p}} \otimes \Gamma_p^\circ)^{1/2} \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & (I_{\tilde{\mathcal{H}}_{\mathbf{S},p}} \otimes \Gamma_p^\circ)^{-1/2} \end{bmatrix}.$$

The condition (5.33) in turn can be rewritten in the form

$$\begin{bmatrix} A^* & C^* \\ B^* & D^* \end{bmatrix} \begin{bmatrix} X & 0 \\ 0 & \bigoplus_{p \in \mathbf{P}} (I_{\tilde{\mathcal{H}}_{\mathbf{R},p}} \otimes \Gamma_p^\circ) \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} - \begin{bmatrix} X & 0 \\ 0 & \bigoplus_{p \in \mathbf{P}} (I_{\tilde{\mathcal{H}}_{\mathbf{S},p}} \otimes \Gamma_p^\circ) \end{bmatrix} \prec 0. \quad (5.34)$$

This last condition (5.34) finally gives us an LMI equivalent to the LOI (5.1) for this case. We note that this analysis is just the discrete-time equivalent of Proposition 8.6 in [22].

Remark 5.7. In our analysis to this point we have considered structure subspaces $\Delta \subset \mathcal{L}(\ell^2 \otimes \mathbb{C}^N)$ defined by spatial constraints (block diagonal matrix representation) without any dynamic constraints. It is natural to impose some additional constraints involving dynamics or parameter restrictions (e.g., forcing the parameters to be real)—see [22, pages 255–256] as well as [35]. In this extended remark we discuss some of these additional considerations which have been discussed in the literature.

Consider the setting of Theorem 5.1 but with the structure subspace $\Delta_{\overline{\mathbf{G}}}$ as in (3.18) or (3.19) (with $\mathcal{K} = \ell^2$) replaced by

$$\Delta_{\overline{\mathbf{G}}, \mathbf{TV}} = \{W = \text{diag}_{k=1, \dots, K} [W_k \otimes I_{\mathcal{H}_{p_k}^\circ}] : W_k \in \mathcal{L}(\ell^2)^{n_k \times m_k} \text{ and } W_k V = V W_k\}, \quad (5.35)$$

i.e., $\Delta_{\overline{\mathbf{G}}, \mathbf{TV}}$ consists of those elements of $\Delta_{\overline{\mathbf{G}}}$ which are also shift-invariant. Then, for the case where M is as in Remark 5.6 (i.e., given via multiplication by a rational transfer function $\widehat{M}(\lambda)$ after transforming to the frequency domain via the Z -transform), one can argue that $\mu_{\Delta_{\overline{\mathbf{G}}, \mathbf{TV}}}(M)$ is given by a supremum of a pointwise structured singular value for the matrix function $\widehat{M}(\zeta)$:

$$\mu_{\Delta_{\overline{\mathbf{G}}, \mathbf{TV}}}(M) = \sup_{\zeta \in \mathbb{T}} \mu_{\Delta_{\overline{\mathbf{G}}}}(\widehat{M}(\zeta)). \quad (5.36)$$

Indeed, this point is argued in detail in [22, Theorem 8.22] for the case where the structure space $\Delta_{\overline{\mathbf{G}}}$ has the special form (1.1) (square blocks with only scalar blocks have higher multiplicity); it is now straightforward to adapt the argument to the more general structure $\Delta_{\overline{\mathbf{G}}}$. As we have already noted in Section 1, computation of $\mu_{\Delta_{\overline{\mathbf{G}}}}(\widehat{M}(\zeta))$ at a fixed value of ζ is problematical, hence computation of the supremum in (5.36) is even more so. A natural upper bound for $\mu_{\Delta_{\overline{\mathbf{G}}, \mathbf{TV}}}(M)$ is the frequency-dependent D -scaling

$$\widehat{\mu}_{\Delta_{\overline{\mathbf{G}}, \mathbf{TV}}}(M) := \inf_D \sup_{\zeta \in \mathbb{T}} \{ \|(I_{\tilde{\mathcal{H}}_{\mathbf{R},p}} \otimes D(\zeta)) \widehat{M}(\zeta) (I_{\tilde{\mathcal{H}}_{\mathbf{R},p}} \otimes D(\zeta))^{-1}\|$$

where the infimum can be taken over $D(\zeta)$ equal to a stable rational matrix function invertible on \mathbb{T} . As discussed in Section 1 above, this upper bound is arbitrarily bad (in various technical senses) when taken at a fixed frequency $\zeta \in \mathbb{T}$, and hence has no chance of being sharp for this frequency-dependent situation.

Poolla-Tikku [38] provide a different perspective on this issue, by giving a robust control interpretation to the quantity $\widehat{\mu}_{\Delta_{\overline{\mathbf{G}}, \mathbf{TV}}}(M)$. Extending the setting of [38] to our set of structured uncertainties $\Delta_{\overline{\mathbf{G}}}$, for a positive parameter ν we let $\Delta_{\overline{\mathbf{G}}, \nu}$ consist of those operators Δ in $\Delta_{\overline{\mathbf{G}}}$ such that

$$\|V\Delta - \Delta V\| \leq \nu$$

where V as usual is the forward shift operator on ℓ^2 (of whatever multiplicity fits the context). Thus operators in $\Delta_{\overline{\mathbf{G}}, \nu}$ are constrained to be *slowly time-varying*,

with precise amount of slowness measured by ν (the smaller the ν the more slow is the time variance with $\nu = 0$ corresponding to time-invariance and $\nu = 2\|\Delta\|$ correspondence to no restriction at all). A corollary of the more precise results from [38], again for the classical spacial case where Δ is given by (1.1), is the following: $\hat{\mu}_{\Delta_{\bar{\mathbf{G}}, \text{TV}}}(M) < 1$ if and only if there is some $\nu > 0$ so that $\mu_{\Delta_{\bar{\mathbf{G}}, \nu}}(M) < 1$. As any disturbance in practice can be expected to have some time-variance, it is argued in [38] that computation of the upper bound $\hat{\mu}_{\Delta_{\bar{\mathbf{G}}, \text{TV}}}(M)$ makes more sense physically than the original quantity $\mu_{\Delta_{\bar{\mathbf{G}}, \text{TV}}}(M)$. Followup work of Paganini [36] (see also Chapter 3 of [35]) showed how one can incorporate time-invariant and time-variant blocks as well as blocks with parametric uncertainty simultaneously. The paper of K ro glu-Scherer [31] refines the results still further for a general block structure (possibly nonsquare blocks with arbitrary multiplicities) with preassigned bounds on the time-variation of the blocks, obtaining upper and lower bounds on the optimal possible performance for this general setting. Much of this work (including the book [22]) also incorporates a causality constraint on the original plant and the admissible perturbation operators Δ . Recent work of Scherer-K se [42] analyzes the application of frequency-dependent D -scaling techniques to the somewhat more general setup of a gain-scheduled feedback configuration.

6. THE ENHANCED UNCERTAINTY STRUCTURE OF BERCOVICI-FOIAS-KHARGONEKAR-TANNENBAUM

An alternative enhancement $\tilde{\mu}_{\Delta}(M)$ of the structured singular value $\mu_{\Delta}(M)$ leading to an equality with the upper bound $\tilde{\mu}_{\Delta}(M) = \hat{\mu}_{\Delta}(M)$ was introduced and developed by Bercovici, Foias and Tannenbaum in [15]. Later work with Khar-gonekar [17, 16] obtained an extension to infinite-dimensional situations. Here we show how the main result can be obtained as a simple adaptation of the convexity-analysis approach of Dullerud-Paganini. The following result is essentially Theorem 3 from [15] with a couple of modifications: our result is more general in that we allow Δ to have nonsquare blocks and hence not a C^* -algebra; on the other hand here we consider only the multiplicity-1 case so we are not allowing the structure Δ to be a general C^* -subalgebra as in [15]. The result can also be seen to follow as a corollary of the more general results concerning robustness with respect to mixed linear-time-varying/linear-time-invariant structured uncertainty (see Chapter 3 of [35]).

The setup is close to that of Theorem 3.2 with a couple of differences. We let \mathbf{G}° be a multiplicity-1 M -graph; thus the spaces $\mathcal{H}_p^\circ = \mathbb{C}$ for all $p \in \mathbf{P}$. We therefore generate the source and range coefficient spaces

$$\mathcal{H}_{\mathbf{S}}^\circ = \bigoplus_{p \in \mathbf{P}} \mathcal{H}_{\mathbf{S},p}^\circ \text{ where } \mathcal{H}_{\mathbf{S},p}^\circ = \bigoplus_{s: [s]=p} \mathbb{C}, \quad \mathcal{H}_{\mathbf{R}}^\circ = \bigoplus_{p \in \mathbf{P}} \mathcal{H}_{\mathbf{R},p}^\circ \text{ where } \mathcal{H}_{\mathbf{R},p}^\circ = \bigoplus_{r: [r]=p} \mathbb{C}$$

and the structure subspace

$$\Delta_{\mathbf{G}^\circ} = \text{diag}_{p \in \mathbf{P}} \mathcal{L}(\mathcal{H}_{\mathbf{R},p}^\circ, \mathcal{H}_{\mathbf{S},p}^\circ)$$

For the enhanced structure we proceed as in Subsection 3.3, but with $\mathcal{K} = \mathcal{H}_{\mathbf{S}}^\circ$. This generates enhanced source and range coefficient spaces

$$\mathcal{H}_{\mathbf{S}} = \mathcal{H}_{\mathbf{S}}^\circ \otimes \mathcal{H}_{\mathbf{S}}^\circ, \quad \mathcal{H}_{\mathbf{S},p} = \mathcal{H}_{\mathbf{S}}^\circ \otimes \mathcal{H}_{\mathbf{S},p}^\circ, \quad \mathcal{H}_{\mathbf{R}} = \mathcal{H}_{\mathbf{R}}^\circ \otimes \mathcal{H}_{\mathbf{R}}^\circ, \quad \mathcal{H}_{\mathbf{R},p} = \mathcal{H}_{\mathbf{R}}^\circ \otimes \mathcal{H}_{\mathbf{R},p}^\circ$$

with the resulting enhanced structure (see (3.18))

$$\Delta_{\overline{\mathbf{G}}} = \text{diag}_{p \in \mathbf{P}} \mathcal{L}(\mathcal{H}_{\mathbf{S}}^{\circ} \otimes \mathcal{H}_{\mathbf{R},p}^{\circ}, \mathcal{H}_{\mathbf{S}}^{\circ} \otimes \mathcal{H}_{\mathbf{S},p}^{\circ}).$$

Then we have the following version of Theorem 3 from [15].

Theorem 6.1. *Let $\Delta_{\mathbf{G}^{\circ}}$ and $\Delta_{\overline{\mathbf{G}}}$ be as above, assume the graph \mathbf{G}° is finite and let M° be any operator from $\mathcal{H}_{\mathbf{S}}^{\circ}$ into $\mathcal{H}_{\mathbf{R}}^{\circ}$. Then*

$$\widehat{\mu}_{\Delta_{\mathbf{G}^{\circ}}}(M^{\circ}) := \mu_{\Delta_{\overline{\mathbf{G}}}}(I_{\mathcal{H}_{\mathbf{S}}} \otimes M^{\circ}) = \widehat{\mu}_{\Delta_{\mathbf{G}^{\circ}}}(M^{\circ}).$$

Proof. We use the identification maps

$$(U_{\mathcal{H}_{\mathbf{S}}, \mathcal{H}_{\mathbf{S}}}^{\circ})^{\top} : \mathcal{H}_{\mathbf{S}}^{\circ} \otimes \mathcal{H}_{\mathbf{S}}^{\circ} \rightarrow \mathcal{C}_2(\mathcal{H}_{\mathbf{S}}^{\circ}), \quad (U_{\mathcal{H}_{\mathbf{S}}, \mathcal{H}_{\mathbf{R}}}^{\circ})^{\top} : \mathcal{H}_{\mathbf{S}}^{\circ} \otimes \mathcal{H}_{\mathbf{S}}^{\circ} \rightarrow \mathcal{C}_2(\mathcal{H}_{\mathbf{S}}^{\circ}, \mathcal{H}_{\mathbf{R}}^{\circ})$$

to view elements of $\mathcal{H}_{\mathbf{S}}$ and $\mathcal{H}_{\mathbf{R}}$ as being in the Hilbert spaces of Hilbert-Schmidt operators $\mathcal{C}_2(\mathcal{H}_{\mathbf{S}}^{\circ})$ and $\mathcal{C}_2(\mathcal{H}_{\mathbf{S}}^{\circ}, \mathcal{H}_{\mathbf{R}}^{\circ})$ from the start. With these identifications, the operator $I \otimes M$ becomes the operator $L_M : \mathcal{C}_2(\mathcal{H}_{\mathbf{S}}^{\circ}) \rightarrow \mathcal{C}_2(\mathcal{H}_{\mathbf{S}}^{\circ}, \mathcal{H}_{\mathbf{R}}^{\circ})$ of left multiplication by M and the structure space becomes

$$\Delta_{\overline{\mathbf{G}}} = \text{diag}_{p \in \mathbf{P}} \mathcal{L}(\mathcal{C}_2(\mathcal{H}_{\mathbf{S}}^{\circ}, \mathcal{H}_{\mathbf{R},p}^{\circ}), \mathcal{C}_2(\mathcal{H}_{\mathbf{S}}^{\circ}, \mathcal{H}_{\mathbf{S},p}^{\circ}))$$

(note that elements of the spaces $\mathcal{L}(\mathcal{C}_2(\mathcal{H}_{\mathbf{S}}^{\circ}, \mathcal{H}_{\mathbf{R},p}^{\circ}), \mathcal{C}_2(\mathcal{H}_{\mathbf{S}}^{\circ}, \mathcal{H}_{\mathbf{S},p}^{\circ}))$ are not required to be left multipliers). For each $p \in \mathbf{P}$ we define maps

$$\phi_p : \mathcal{H}_{\mathbf{R},p} := \mathcal{C}^2(\mathcal{H}_{\mathbf{S}}^{\circ}, \mathcal{H}_{\mathbf{R},p}^{\circ}) \rightarrow \mathbb{C}$$

by

$$\phi_p : h \mapsto \text{tr} \left((M^{\circ*} P_{\mathcal{H}_{\mathbf{R},p}^{\circ}} M^{\circ} - P_{\mathcal{H}_{\mathbf{S},p}^{\circ}}) h h^* \right). \quad (6.1)$$

We set

$$\nabla = \{(\phi_p(h))_{p \in \mathbf{P}} : \|h\|_{\mathcal{H}_{\mathbf{S}}} = 1\}, \quad \Pi = \{(r_p)_{p \in \mathbf{P}} : r_p \in \mathbb{R} \text{ with } r_p \geq 0\}. \quad (6.2)$$

Then we have the following analogue of Lemma 5.2.

Lemma 6.2. *Assume that $\mu_{\Delta_{\overline{\mathbf{G}}}}(L_M) < 1$. Then, with ∇ and Π as in (6.2),*

$$\nabla \cap \Pi = \emptyset.$$

Proof. We proceed by contradiction. Suppose that there is an $h \in \mathcal{H}_{\mathbf{S}} = \mathcal{C}_2(\mathcal{H}_{\mathbf{S}}^{\circ})$ with norm 1 and $\phi_p(h) > 0$ for all p . This means that

$$\|P_{\mathcal{H}_{\mathbf{R},p}^{\circ}} M^{\circ} h\|_{\mathcal{C}_2(\mathcal{H}_{\mathbf{S}}^{\circ}, \mathcal{H}_{\mathbf{R},p}^{\circ})}^2 - \|P_{\mathcal{H}_{\mathbf{S},p}^{\circ}} h\|_{\mathcal{C}_2(\mathcal{H}_{\mathbf{S}}^{\circ}, \mathcal{H}_{\mathbf{S},p}^{\circ})}^2 \geq 0$$

for all $p \in \mathbf{P}$. Here it is understood that the projections and M° act via left multiplication. By the standard Douglas lemma [20], there is a contraction operator Δ_p from $\mathcal{C}_2(\mathcal{H}_{\mathbf{S}}^{\circ}, \mathcal{H}_{\mathbf{R},p}^{\circ})$ to $\mathcal{C}_2(\mathcal{H}_{\mathbf{S}}^{\circ}, \mathcal{H}_{\mathbf{S},p}^{\circ})$ (not necessarily a left multiplier) so that $\Delta_p[P_{\mathcal{H}_{\mathbf{R},p}^{\circ}} M^{\circ} h] = P_{\mathcal{H}_{\mathbf{S},p}^{\circ}} h$. If we set $\Delta = \text{diag}_{p \in \mathbf{P}} \Delta_p$, then Δ is in $\overline{\mathcal{B}}\Delta_{\overline{\mathbf{G}}}$ and $(I - \Delta L_M)h = 0$. Hence $I - \Delta L_M$ is not invertible, contrary to the assumption that $\mu_{\Delta_{\overline{\mathbf{G}}}}(L_M) < 1$. This completes the proof of the lemma. \square

The next lemma (the analogue of Lemma 5.3) gives the connection between $\widehat{\mu}_{\Delta_{\mathbf{G}^{\circ}}}(M^{\circ}) < 1$ and the quadratic forms ϕ_p .

Lemma 6.3. *The condition $\widehat{\mu}_{\Delta_{\mathbf{G}^{\circ}}}(M^{\circ}) < 1$ holds if and only if either of the following conditions holds:*

(1) There are positive numbers r_p ($p \in \mathbf{P}$) so that

$$\sum_{p \in \mathbf{P}} \gamma_p \phi_p(h) \leq -\epsilon \|h\|_{\mathcal{C}_2(\mathcal{H}_{\mathbf{S}}^\circ)}^2$$

for all $h \in \mathcal{C}_2(\mathcal{H}_{\mathbf{S}}^\circ)$.

(2) The sets ∇ and Π (see (6.2)) are strictly separated: there exist real numbers $\gamma_p \in \mathbb{R}$ and numbers $\alpha < \beta$ so that

$$\sum_{p \in \mathbf{P}} \operatorname{Re} \gamma_p k_p \leq \alpha < \beta \leq \sum_{p \in \mathbf{P}} \operatorname{Re} \gamma_p r_p \text{ for } (k_p)_{p \in \mathbf{P}} \in \nabla \text{ and } (r_p)_{p \in \mathbf{P}} \in \Pi.$$

Proof. The condition $\widehat{\mu}_{\Delta_{\mathbf{G}^\circ}}(M^\circ)$ can be expressed as: there exist numbers $\gamma_p > 0$ ($p \in \mathbf{P}$) so that

$$M^* \left(\bigoplus_{p \in \mathbf{P}} \gamma_p I_{\mathcal{H}_{\mathbf{R},p}^\circ} \right) M - \left(\bigoplus_{p \in \mathbf{P}} \gamma_p I_{\mathcal{H}_{\mathbf{S},p}} \right) \prec 0.$$

We rewrite this as the higher multiplicity quadratic-form condition

$$\sum_{p \in \mathbf{P}} \gamma_p \left(\|P_{\mathcal{H}_{\mathbf{R},p}^\circ} M h\|_{\mathcal{C}_2(\mathcal{H}_{\mathbf{S}}^\circ, \mathcal{H}_{\mathbf{R},p}^\circ)}^2 - \|P_{\mathcal{H}_{\mathbf{S},p}} h\|_{\mathcal{C}_2(\mathcal{H}_{\mathbf{S}}^\circ, \mathcal{H}_{\mathbf{S},p})}^2 \right) < -\epsilon^2 \|h\|^2$$

for some $\epsilon > 0$. This can be manipulated to the equivalent form

$$\sum_{p \in \mathbf{P}} \gamma_p \operatorname{tr} \left((M^* P_{\mathcal{H}_{\mathbf{R},p}^\circ} M - P_{\mathcal{H}_{\mathbf{S},p}}) h h^* \right) = \sum_{p \in \mathbf{P}} \gamma_p \phi_p(h) < -\epsilon^2 \|h\|^2$$

verifying (1). The equivalence of (1) and (2) proceeds just as in the proof of Lemma 5.3. \square

To complete the proof, following the same strategy as in the proof of Theorem 5.1, a Hahn-Banach separation theorem (specifically, Theorem 3.4 part (b) in [39]) enables to complete the proof if we can show the strengthened version of the result of Lemma 6.2, namely:

$$\mu_{\Delta_{\overline{\mathbf{G}}}}(L_{M^\circ}) < 1 \quad \Rightarrow \quad \overline{\operatorname{co}} \nabla \cap \Pi = \emptyset. \quad (6.3)$$

In the present setting, the space $\mathcal{H}_{\mathbf{S}} = \mathcal{C}_2(\mathcal{H}_{\mathbf{S}}^\circ)$ is finite-dimensional and hence has compact unit ball. A standard continuity argument then implies that $\nabla = \overline{\nabla}$ is in fact a closed subset in $\mathbb{R}^{\mathbf{P}}$. Thus the only remaining piece to show is that ∇ itself is already convex. This follows from the following lemma.

Lemma 6.4. *The set ∇ given by (6.2) is convex.*

Proof. Suppose that h and \tilde{h} are two unit-norm elements of $\mathcal{C}_2(\mathcal{H}_{\mathbf{S}}^\circ)$ and $0 < \alpha < 1$. We must find a unit-norm $k \in \mathcal{C}_2(\mathcal{H}_{\mathbf{S}}^\circ)$ so that

$$\phi_p(k) = \alpha \phi_p(h) + (1 - \alpha) \phi_p(\tilde{h}) \text{ for all } p \in \mathbf{P}.$$

Towards this end we observe that

$$\alpha \phi_p(h) + (1 - \alpha) \phi_p(\tilde{h}) = \operatorname{tr} \left((M^{\circ*} P_{\mathcal{H}_{\mathbf{R},p}^\circ} M^\circ - P_{\mathcal{H}_{\mathbf{S},p}}) (\alpha h h^* + (1 - \alpha) \tilde{h} \tilde{h}^*) \right). \quad (6.4)$$

Note that the operator $\Upsilon := \alpha h h^* + (1 - \alpha) \tilde{h} \tilde{h}^*$ is a trace class operator of rank at most $\dim \mathcal{H}_{\mathbf{S}}^\circ$. Therefore we may factor Υ in the form $\Upsilon = k k^*$ where $k \in \mathcal{C}_2(\mathcal{H}_{\mathbf{S}}^\circ)$.

Furthermore

$$\begin{aligned}\|k\|_{\mathcal{C}_2(\mathcal{H}_{\mathbb{S}}^{\circ})}^2 &= \operatorname{tr}(k^*k) = \operatorname{tr}(\alpha h h^* + (1-\alpha)\widetilde{h}\widetilde{h}^*) = \alpha \operatorname{tr}(h h^*) + (1-\alpha)\operatorname{tr}(\widetilde{h}\widetilde{h}^*) \\ &= \alpha \|h\|^2 + (1-\alpha)\|\widetilde{h}\|^2 = 1\end{aligned}$$

so k also has unit norm. Finally, from (6.4) we read off

$$\phi_p(k) = \operatorname{tr}\left((M^{\circ*}P_{\mathcal{H}_{\mathbb{R},p}^{\circ}}M^{\circ} - P_{\mathcal{H}_{\mathbb{S},p}^{\circ}})kk^*\right) = \alpha\phi_p(h) + (1-\alpha)\phi_p(\widetilde{h})$$

as wanted. This completes the proof of the lemma. \square

Putting all these pieces together completes the proof of Theorem 6.1. \square

REFERENCES

- [1] J. Agler, On the representation of certain holomorphic functions defined on a polydisk, in: *Topics in Operator Theory: Ernst D. Hellinger Memorial Volume*, pp. 47–66, Oper. Theory Adv. Appl. **48**, Birkhäuser-Verlag, Basel, 1990.
- [2] J. Agler and J.E. McCarthy, Global holomorphic functions in several non-commuting variables, arXiv:1305.1636v1, 2013.
- [3] J. Agler and J.E. McCarthy, Pick interpolation for free holomorphic functions, arXiv:1403.1960.
- [4] D. Alpay and D.S. Kalyuzhnyi-Verbovetzkiĭ, Matrix- J -unitary non-commutative rational formal power series, in: *The State-Space Method: Generalizations and Applications*, pp. 49–113, Oper. Theory Adv. Appl. **161**, Birkhäuser, Basel, 2006.
- [5] C.-G. Ambrozie and D. Timotin, A von Neumann type inequality for certain domains in \mathbb{C}^n , *Proc. Amer. Math. Soc.* **131** (2003), 859–869.
- [6] B.D.O. Anderson, P. Agathoklis, E.I. Jury, and M. Mansour, Stability and the matrix Lyapunov equation for discrete 2-dimensional systems, *IEEE Trans. Circuits & Systems* **33** (1986), 261–267.
- [7] D.Z. Arov, M.A. Kaashoek, and D.R. Pik, The Kalman-Yakubovich-Popov inequality for discrete time systems of infinite dimension, *J. Operator Theory* **55** (2006), 393–438.
- [8] J.A. Ball and V. Bolotnikov, Realization and interpolation for Schur-Agler-class functions on domains with matrix polynomial defining function in \mathbb{C}^n , *J. Funct. Anal.* **213** (2004), 45–87.
- [9] J.A. Ball, G. Groenewald, and T. Malakorn, Structured noncommutative multidimensional linear systems, *SIAM J. Control Optim.* **44** (2005), 1474–1528.
- [10] J.A. Ball, G. Groenewald, and T. Malakorn, Conservative structured noncommutative multidimensional linear systems, in: *The State Space Method: Generalizations and Applications* pp. 179–223, Oper. Theory Adv. Appl. **161**, Birkhäuser, Basel, 2006.
- [11] J.A. Ball, G. Groenewald, and T. Malakorn, Bounded real lemma for structured noncommutative multidimensional linear systems and robust control, *Multidimens. Syst. Signal Process.* **17** (2006), 119–150.
- [12] J.A. Ball, G. Groenewald, and S. ter Horst, Structured singular values versus diagonal scaling: the noncommutative setting, Proceedings of MTNS 2014, Groningen, to appear.
- [13] J.A. Ball and T. Malakorn, Multidimensional linear feedback control systems and interpolation problems for multivariable holomorphic functions, *Multidimens. Systems Signal Process.* **15** (2004), 7–36.
- [14] J.A. Ball and S. ter Horst, Robust control, multidimensional systems and multivariable Nevanlinna-Pick interpolation, in: *Topics in Operator Theory Volume 2: Systems and Mathematical Physics*, pp. 13–88, Oper. Theory Adv. Appl. **161**, Birkhäuser, Basel, 2010.
- [15] H. Bercovici, C. Foias and A. Tannenbaum, Structured Interpolation Theory, in: *Extension and Interpolation of Linear Operators and Matrix Functions*, pp. 195–220, Oper. Theory Adv. Appl. **203**, Birkhäuser, Basel, 1990.
- [16] H. Bercovici, C. Foias and A. Tannenbaum, The structured singular value for linear input/output operators, *SIAM J. Control and Optimization* **34** (1996), 1392–1404.
- [17] H. Bercovici, C. Foias, P.P. Khargonekar, and A. Tannenbaum, On a lifting theorem for the structured singular value, *J. Math. Anal. Appl.* **187** (1994) 617–627.

- [18] R. Braatz, P. Young, J.C. Doyle, and M. Morari, Computational complexity of μ calculations, *IEEE Transactions on Automatic Control* **39** (1994), 1000–1002.
- [19] J. Cimprič, J.W. Helton, I. Klep, S. McCullough, and C. Nelson, On real one-sided ideals in a free algebra, *J. Pure Appl. Algebra* **218** (2014), 269–284.
- [20] R.G. Douglas, On majorization, factorization, and range inclusion of operators on Hilbert space, *Proc. Amer. Math. Soc.* **17** (1966), 413–415.
- [21] J.C. Doyle, Analysis of feedback systems with structured uncertainties, *IEEE Proceedings* **129** (1982), 242–250.
- [22] G.E. Dullerud and F. Paganini, *A Course in Robust Control Theory: A Convex Approach*, Texts in Applied Mathematics Vol. **36**, Springer-Verlag, New York, 2000.
- [23] M. Fliess, Matrices de Hankel, *J. Math. Pures Appl.* **53** (1974), 197–222; erratum, **54** (1975), 481.
- [24] J.W. Helton, Positive noncommutative polynomials are sums of squares, *Ann. of Math.* **156** (2002), 675–694.
- [25] J.W. Helton, I. Klep, S. McCullough, and N. Slingsend, Noncommutative ball maps, *J. Funct. Anal.* **257** (2009), 47–87.
- [26] J.W. Helton, I. Klep, and S. McCullough, Analytic mappings between noncommutative pencil balls, *J. Math. Anal. Appl.* **376** (2011), 407–428.
- [27] J.W. Helton and S. McCullough, Every convex free basic semi-algebraic set has an LMI representation, *Ann. of Math.* **356** (2012), 979–1013.
- [28] J.W. Helton and S. McCullough, A Positivstellensatz for non-commutative polynomials, *Trans. Amer. Math. Soc.* **356** (2004), 3721–3737.
- [29] J.W. Helton, S. McCullough, and V. Vinnikov, Noncommutative convexity arises from linear matrix inequalities, *J. Funct. Anal.* **240** (2012), 105–191.
- [30] D.S. Kaliuzhnyi-Verbovetskyi and V. Vinnikov, *Foundations of Noncommutative Function Theory*, arXiv:1212:6345.
- [31] H. Köroğlu and C.W. Scherer, Robust performance analysis for structured linear time-varying perturbations with bounded rates-of-variation, *IEEE Trans. Automat. Control* **52** (2007) No. 2, 197–211.
- [32] W.-M. Lu, K. Zhou, and J.C. Doyle, Stabilization of uncertain linear systems: an LFT approach, *IEEE Trans. Auto. Control* **41** (1998), 50–65.
- [33] A. Megretski and S. Treil, Power distribution inequalities in optimization and robustness of uncertain systems, *J. Mathematical Systems, Estimation, and Control* **3** (1993), 301–319.
- [34] A. Packard and J.C. Doyle, The complex structured singular value, *Automatica* **29** (1993), 71–109.
- [35] F. Paganini, *Sets and Constraints in the Analysis of Uncertain Systems*, Thesis submitted to California Institute of Technology, Pasadena, 1996.
- [36] F. Paganini, Robust stability under mixed time-varying, time-invariant and parametric uncertainty, *Automatica* **32** (1996) NO. 10, 1381–1392.
- [37] I.R. Petersen, B.D.O. Anderson, and E.A. Jonckheere, *A first principles solution to the non-singular H^∞ control problem*, Internat. J. Robust Nonlinear Control **1** (1991), 171–185.
- [38] K. Poolla and A. Tikku, Robust performance against time-varying structured perturbations, *IEEE Trans. Automat. Control* **40** (1995), 1589–1602.
- [39] W. Rudin, *functional Analysis*, McGraw-Hill, New York, 1973; Second Edition, 1991.
- [40] M.G. Safanov, *Stability Robustness of Multivariable Feedback Systems*, MIT Press, Cambridge, MA, 1980.
- [41] J.S. Shamma, Robust stability with time-varying structured uncertainty, *IEEE Trans. Automat. Control* **39** (1994), 714–724.
- [42] C.W. Scherer and I.E. Köse, Gain-scheduled control synthesis using dynamic D -scales, *IEEE Trans. Automat. Control* **57** (2012) No. 9, 714–724.
- [43] M. Takesaki, *Theory of Operator Algebras I*, Encyclopedia of Mathematical Sciences **124**, Springer, New York (1979).
- [44] S. Treil, The gap between the complex structured singular value μ and its upper bound is infinite, preprint.
- [45] K. Zhou, J.C. Doyle and K. Glover, *Robust and Optimal Control*, Prentice-Hall, Upper Saddle River, NJ, 1996.

DEPARTMENT OF MATHEMATICS, VIRGINIA TECH, BLACKSBURG, VA 24061-0123, USA
E-mail address: `joball@math.vt.edu`

DEPARTMENT OF MATHEMATICS, UNIT FOR BMI, NORTH-WEST UNIVERSITY, POTCHEFSTROOM,
2531 SOUTH AFRICA
E-mail address: `Gilbert.Groenewald@nwu.ac.za`

DEPARTMENT OF MATHEMATICS, UNIT FOR BMI, NORTH-WEST UNIVERSITY, POTCHEFSTROOM,
2531 SOUTH AFRICA
E-mail address: `Sanne.TerHorst@nwu.ac.za`